

Université de Limoges

ED 653 – Sciences et Ingénierie

Faculté des Sciences et Techniques – Institut de Recherche XLIM

**Thèse pour obtenir le grade de
Docteur de l'Université de Limoges
Électronique des Hautes Fréquences, Photonique et Systèmes**

Présentée et soutenue par

Soukayna EL KARAKHI

Le 11 Décembre 2023

**APPROCHES AVANCÉES EN APPRENTISSAGE PROFOND POUR LE DIAGNOSTIC DES
CÂBLES ET L'ANALYSE D'IMAGES RADAR GPR**

Thèse dirigée par Alain REINEIX et Christophe GUIFFAUT

JURY :

Examineur

Mme. Sonia Bendhia, Professeur des universités, LAAS-CNRS, INSA Toulouse

Rapporteurs

M. Boyer Alexandre, Maître de conférences, LAAS-CNRS, INSA Toulouse

M. Lionel Pichon, Directeur de Recherche C.N.R.S, Université Paris-Saclay, rattachement

Directeur de thèse

M. Alain Reineix, Directeur de Recherche C.N.R.S, XLIM UMR CNRS 7252, Université de Limoges

Co-directeur de thèse

M. Guiffaut Christophe, Maître de conférences, XLIM UMR CNRS 7252, Université de Limoges



Resumé

Titre: Approches Avancées en Apprentissage Profond pour le diagnostic des câbles et l'analyse d'images Radar GPR.

Mot clés: Diagnostic filaire complexe; Défaut non franc; Réflectométrie dans le domaine temporel; Réflectométrie distribuée; Radar à Pénétration de Sol; Apprentissage profond; Réseaux de neurones convolutifs.

Au cours des dernières années, les méthodes d'apprentissage profond, ou "Deep Learning", ont significativement évolué et trouvé des applications dans divers domaines. Ce travail de thèse explore son applicabilité dans deux domaines spécifiques de la physique : la réflectométrie pour l'identification de défauts le long des réseaux de câblages, et la détection d'objets enterrés à l'aide du radar à sondage de sol (GPR) grâce à des techniques non destructives. Les câbles, éléments critiques dans les systèmes électroniques, peuvent présenter des défaillances, représentant ainsi un risque potentiel dans des secteurs sensibles comme l'aéronautique. L'utilisation de la réflectométrie, une technique exploitant les signaux réfléchis après envoi d'un signal dans le câble, est vitale pour détecter et gérer des anomalies naissantes. L'intégration de réseaux de neurones profonds, incluant des modèles à multi-perceptrons et convolutionnels, a démontré son efficacité pour localiser et caractériser précisément les défauts. Au cours de cette thèse, nous avons mis en lumière l'influence des hyperparamètres sur la précision du modèle construit. Dans le domaine des radars GPR, une antenne émettrice génère des impulsions électromagnétiques pour illuminer des cibles situées dans le sous-sol. Le traitement des signaux réfléchis lors du déplacement du récepteur permet de construire une image, appelée B-scan, qui affiche des hyperboles fournissant des informations sur la localisation et la nature du tuyau. Grâce à la mise en œuvre de réseaux neuronaux profonds tels que VGG-16, ResNet-50, et des CNN, adaptés à partir d'une base de données synthétiques, un modèle de classification multi-labels a été développé. Celui-ci permet une extraction instantanée des caractéristiques du sol et des tuyaux (diamètre, type de tuyau), améliorant significativement l'efficacité de la détection et de l'analyse. Un réglage rigoureux des hyperparamètres s'est avéré nécessaire.

Abstract

Title: Advanced Approaches in Deep Learning for Cable Diagnostics and GPR Radar Image Analysis.

Keywords: Complex Wireline Diagnosis; Non-clear Fault; Time-domain Reflectometry; Distributed Reflectometry; Ground Penetrating Radar; Deep Learning; Convolutional Neural Networks.

In recent years, deep learning methods have significantly evolved and found applications in various fields. This thesis work explores its applicability in two specific physics domains: reflectometry for identifying faults along cable networks, and the detection of buried objects using Ground Penetrating Radar (GPR) through non-destructive techniques. Cables, being critical components in electronic systems, can exhibit failures, thus posing a potential risk in sensitive sectors like aeronautics. The use of reflectometry, a technique that employs signals reflected back after being sent through the cable, is crucial for identifying and addressing emerging anomalies. The integration of deep neural networks, including multi-perceptron and convolutional models, has proven effective for precisely locating and characterizing defects. During this thesis, we highlighted the influence of hyperparameters on the accuracy of the constructed model. In GPR applications, an emitting antenna discharges electromagnetic pulses to spotlight underground targets. Processing the signals reflected during the receiver's movement enables the formation of a B-scan image, showcasing hyperboles that convey information about the position and type of pipe. Using deep neural networks, specifically VGG-16, ResNet-50, and customized CNNs, derived from a synthetic database, a multi-label classification model was formulated. This model facilitates the instantaneous extraction of soil and pipe attributes (diameter, type of pipe), markedly improving detection and analysis efficacy. A rigorous adjustment of the hyperparameters proved to be indispensable.

Dédicace

Vous pouvez inscrire ici une citation, par exemple :

If I have seen further, it is by standing on the shoulders of Giants.

Sir Isaac Newton

Remerciements

Écrivez vos remerciements dans cette rubrique (professeurs, amis, famille, etc).

Droits d'auteurs

Cette création est mise à disposition selon le Contrat :

« Attribution-Pas d'Utilisation Commerciale-Pas de modification 3.0 France »

disponible en ligne : <http://creativecommons.org/licenses/by-nc-nd/3.0/fr/>



Table des Matières

1 Introduction a l'apprentissage automatique et état de l'art sur l'apprentissage profond	22
1.1 Introduction	23
1.2 Apprentissage automatique	24
1.2.1 Expérience E et tâche T	25
1.2.2 Apprentissage supervisé	26
1.2.3 Classification	26
1.2.4 Régression	29
1.2.5 Apprentissage non supervisé	30
1.2.6 Performance P	31
1.3 Introduction aux Réseaux de neurones artificiels	34
1.3.1 Retour sur les réseaux de neurones humains	34
1.3.2 Neurones artificiels (neurone formel)	35
1.3.3 Historique et évolution des réseaux de neurones	37
1.4 Apprentissage des réseaux de neurones artificiels	38
1.4.1 Définition d'un perceptron multicouche (MLP)	38
1.4.2 Algorithme de descente de gradient	39
1.4.3 Types de fonctions d'activation	41
1.4.4 Paramétrage d'un modèle d'apprentissage profond	44
1.5 Réseau de neurones convolutifs	47
1.5.1 Architecture de base des Réseau de neurones convolutifs	48
1.5.2 Couche de convolution	49
1.5.3 Type de convolution	52
1.5.4 Couche de pooling	54
1.5.5 Couche complètement connectée	56
1.5.6 Paramètres spécifiques aux réseaux convolutifs	57
1.6 Architectures avancées	60
1.6.1 VGG	60

1.6.2	Resnet50	62
1.7	Conclusion	63
2	Etat de l'art sur la détection et la localisation de défauts dans les câbles par réflectométrie	64
2.1	Introduction	65
2.2	Domaines d'application des câbles	66
2.3	Types de câbles électriques	68
2.3.1	Le câble coaxial	68
2.3.2	Paire torsadé	69
2.3.3	Câbles d'énergie	71
2.4	Classification des Anomalies Présentes dans les câbles électriques . .	72
2.4.1	Défauts non francs	72
2.4.2	Défauts sévères	74
2.5	Etat de l'art sur le diagnostic des défauts dans les câbles	75
2.6	Concept de réflectométrie et théorie des lignes	76
2.7	Modélisation d'une ligne	78
2.7.1	Régime temporel quelconque	78
2.7.2	Solutions générales en régime sinusoïdal	79
2.7.3	Coefficient de réflexion	81
2.7.4	Modélisation de défauts	83
2.8	Méthodes basées sur la réflectométrie	84
2.8.1	Réflectométrie dans le domaine temporel	84
2.9	Réflectométrie classique	84
2.9.1	STDR/SSTDR	85
2.9.2	Réflectométrie dans le domaine fréquentiel	86
2.10	Conclusion	89
3	Diagnostic de l'état d'un réseau filaire complexe	91
3.1	Introduction	92
3.2	Topologie de câbles complexe	92
3.2.1	Ambiguïté de localisation des défauts	94
3.3	Description de la base de données	95
3.4	Mise en oeuvre des réseaux de neurones classiques pour la détection des défauts non francs	98
3.4.1	Modèle de régression	98
3.4.2	Influence des hyperparamètres sur le réseau MLP	103

3.4.3	Influence de fonction d'activation	104
3.4.4	Influence des optimiseurs	104
3.4.5	Influence de nombre de neurones	105
3.5	Mise en œuvre des réseaux de neurones convolutifs pour la détection des défauts non francs	107
3.6	Sélection des hyper-paramètres optimaux	107
3.6.1	Influence du nombre des couches de convolution	107
3.6.2	Influence du nombre de cachées sur les performances	108
3.6.3	Influence de la taille des filtres de convolution	109
3.7	Synthèse	111

4 Principes Fondamentaux de l'Électromagnétisme Appliqués aux Radars

GPR		113
4.1	Introduction	115
4.2	Notions d'électromagnétisme appliquée au radar	115
4.2.1	Les équations de propagation des ondes électromagnétiques dans les milieux	115
4.3	Comportement de la matière vis à vis du champ magnétique	119
4.4	Comportement de la matière vis à vis du champ électrique	120
4.5	Propriétés caractéristiques des sols	122
4.6	Principe de fonctionnement et caractéristiques d'un radar GPR	123
4.6.1	Historique et domaines d'applications du GPR	124
4.6.2	Différents types de radar	125
4.6.3	Radar à modulation linéaire de fréquence (Frequency Modulated Continuous Wave)	126
4.6.4	Autres types de radar	127
4.6.5	Le radar RAMAC	127
4.7	Performance du radar GPR et Imagerie Radar	128
4.7.1	Équation du radar	128
4.7.2	Dynamique du système	129
4.7.3	Profondeur d'investigation	130
4.7.4	Résolution en profondeur	130
4.8	Les enregistrement radar	130
4.8.1	A-Scan	131
4.8.2	B-Scan	132
4.8.3	C-Scan	132

4.9	Les différents modes d'acquisition	133
4.9.1	Mode réflexion	133
4.10	Antennes GPR	135
4.10.1	Directivité et Gain	135
4.10.2	Impédance d'entrée de l'antenne	136
4.10.3	Bande passante	137
4.10.4	Polarisation d'une antenne	138
4.10.5	Principaux types d'antennes GPR	139
4.11	Méthodes numériques pour la construction de B-scan	141
4.11.1	Choix de la méthode d'analyse numérique	141
4.11.2	Principe de la méthode FDTD	143
4.11.3	Stabilité de calcul et dispersion numérique	146
4.11.4	Modélisation de l'espace infini	147
4.11.5	Modélisation des antennes filaires	147
4.12	Conclusion	148
5	Contribution et développement des architectures de l'apprentissage profond pour la classification des images B-scan	149
5.1	Introduction	151
5.2	Création de la base de données	151
5.2.1	Configuration Multi-Capteurs	151
5.2.2	Antennes Wu-King	152
5.2.3	Signal d'excitation	154
5.2.4	Milieu de propagation	155
5.3	Modélisation FDTD pour la création de la base de données	157
5.4	Tracé des B-scans et recalage temporel	158
5.5	Prétraitement des données	161
5.5.1	Labellisation des images B-scan	162
5.6	Détection des contours	163
5.6.1	Principe d'opérateur de gradient	163
5.6.2	Sobel	164
5.6.3	Canny	165
5.6.4	Augmentation des données (data augmentation)	167
5.6.5	Séparation des données	167
5.6.6	Encodage des étiquettes	168
5.7	Architectures utilisées pour la multi label classification des images B-Scans	169

5.7.1	Principe de transfert par apprentissage (transfert Learning)	169
5.7.2	Architecture Resnet 50	170
5.8	Architecture VGG16	177
5.9	CNN personnalisé	180
5.10	Influence des hyperparametres	182
5.10.1	Méthodologie	183
5.11	Synthèse	189
	Références	195

Table des Figures

1.1	Diagramme de Venn représentant les relations des domaines liées à l'IA	24
1.2	Méthodes d'Apprentissage Automatique	25
1.3	Classification binaire	27
1.4	Classification binaire	31
1.5	Analogie entre le neurone biologique et neurone artificiel [17]	35
1.6	Fonctionnement d'un Neurone Artificiel	36
1.7	Perceptron multicouche (MLP)	38
1.8	Illustration des fonctions d'activation bornées et non bornées.	42
1.9	Fonctions d'activation selon la tâche : Classification vs Régression	44
1.10	Capacité, surajustement et sous-ajustement	45
1.11	Relation typique entre l'erreur de prédiction et la complexité du modèle	46
1.12	Illustration de la technique de Dropout : désactivation aléatoire de neurones	47
1.13	Illustration d'une architecture de réseau de type CNN.	49
1.14	Convolution 2D et Génération de Carte Caractéristique	50
1.15	Convolution 3D	52
1.16	Opération de convolution [31]	53
1.17	À gauche, exemple de max-pooling et à droite, exemple de average-pooling, d'une carte d'activation de taille 7×7 avec des strides de 1 et 2.	56
1.18	Couche complètement connecté	57
2.1	Longueurs cumulées de câbles dans les différentes applications (en km)	65
2.2	Description matérielle d'un câble coaxial (source: Wikipédia).	69
2.3	Figure montrant les trois types de paires torsadées: non blindées, blindées et super blindées (source : Wikipédia).	71
2.4	Coupe transversale d'un câble d'énergie (source wikipedia)	72
2.5	Types des défauts et leur répartitions selon la base de données Navy Safety Center Hazardous Incident entre 1980 et 1999 [57].	73

2.6	Exemples de défauts: (a) circuit-ouvert (défaut franc), (b) perte de l'isolant (défaut non franc) et (c) défaut d'écrasement (défaut non franc) [58]. . .	74
2.7	Exemple de défauts sévères dans les câbles aéronautiques [59].	75
2.8	Principe de la réflectométrie	77
2.9	Réfectogrammes pour différentes valeurs de Z_l sur un câble coaxial d'impédance caractéristique $Z_c = 50 \Omega$	78
2.10	Circuit électrique équivalent à un tronçon de ligne de transmission, montrant les paramètres linéiques R, L, C et G	79
2.11	Modèle d'une ligne de transmission de longueur L et d'impédance caractéristique Z_t	81
2.12	Modélisation du défaut distribué	83
2.13	Principe de réflectométrie classique	85
2.14	Exemple de signal chirp linéaire avec son spectre fréquentiel couvrant une bande passant de 40MHz.	87
3.1	Réseau bifilaire à 5 branches et deux défauts non francs sur la branche B1 et la branche B3	92
3.2	Réfectogramme au niveau de R1	93
3.3	Réfectogramme au niveau de R2	94
3.4	Les deux pics du défaut non franc, positionné aléatoirement et détecté par R2 sur la branche B3.	96
3.5	La forme d'onde du signal réfléchi par le défaut non franc : a) à gauche, le pic positif à l'entrée du toron défectueux; b) à droite, le pic négatif à la sortie du toron défectueux	97
3.6	Schéma RN	99
3.7	Attributs avant et après la normalisation	100
3.8	Score MAE pour l'ensemble de training vs validation	103
3.9	Score MSE pour l'ensemble de training vs validation	103
3.10	Influence de la fonction d'activation	104
3.11	Influence des optimiseurs	105
3.12	Influence du nombre de neurones	106
3.13	Impact du nombre de couches de convolution	108
3.14	Impact du nombre de couches cachées	109
3.15	Impact de la taille des filtres de convolution	110
4.1	Type de polarisation en fonction de la fréquence	121
4.2	Éléments composants un GPR	123

4.3	Forme d'onde Chirp	126
4.4	Radar GPR RAMAC	128
4.5	Exemple de la forme d'onde d'un A-scan capté par une antenne réceptrice	131
4.6	Visualisation d'un radagramme: B-Scan.	132
4.7	Représentation du type Cscan	133
4.8	Mode avec offset constant : Acquisition en déplaçant les antennes le long d'une ligne horizontale	134
4.9	Mode avec offset variable	135
4.10	Schéma de l'impédance d'entrée de l'antenne Z_e avec alimentation par un générateur	137
4.11	Illustration des trois types de polarisation : linéaire, circulaire, et elliptique.	138
4.12	Cornet rigide large bande	140
4.13	Antennes papillons.	141
4.14	Antenne wu king et répartition des charges résistives.	141
4.15	Schéma d'une cellule de Yee en 3D.	144
4.16	Schéma saute-mouton (leapfrog) représentant le principe de mise à jour des champs électrique et magnétique	145
5.1	GPR multi-capteur	152
5.2	Réponse temporelle en courant avec un générateur de tension délivrant une forme d'onde sino-gaussienne	153
5.3	Scène GPR avec un tuyau métallique de diamètre 32mm enfoui dans un sol d'argile sèche	158
5.4	La création de l'image B-scan par la juxtaposition de tous les signaux consécutifs A-scan, après soustraction, pour chaque A-scan, des signaux captés sur la même scène sans objet diffractant	160
5.5	Schéma illustrant le retard entre l'impulsion émise et reçue due à la diffraction du tuyau.	161
5.6	Exemples de radagrammes.	163
5.7	166
5.8	Exemples d'augmentation des données radagrammes	167
5.9	Illustrations de schémas de partitionnement des données d'entraînement, validation et de test.	168
5.10	Stratégie globale de transfert par apprentissage.	170
5.11	Architecture Resnet-50	171
5.12	Multi-branches entièrement connectées	172

5.13	Exactitude de la validation pour ResNet-50 au fil des époques.	173
5.14	Perte de la validation pour ResNet-50 au fil des époques.	173
5.15	Perte d'entraînement pour ResNet-50 au fil des époques.	174
5.16	Matrice de confusion pour l'architecture ResNet-50.	175
5.17	Stratégie globale de transfert par apprentissage.	178
5.18	Les résultats de VGG-16 représentant l'évolution de l'entraînement et de la validation (a) de la perte et (b) de l'exactitude au fil des époques. . . .	179
5.19	Architecture CNN de référence	182
5.20	Résultats historiques pour les pertes de validation et d'entraînement. . .	186
5.21	Résultats historiques pour la précision de validation et d'entraînement. .	186
5.22	Comparaison des scores moyens des trois architectures	190

Liste des Tableaux

3.1	Métriques sur l'ensemble de base d'entraînement et de test	102
3.2	Test Data	103
3.3	Meilleurs hyperparamètres pour le réseau MLP.	106
3.4	Paramètres optimaux pour Rank 1 et Rank 2 après optimisation	111
4.1	Caractéristiques EM des matériaux mesurées à 100 MHz.	122
5.1	Tableau des valeurs de résistance sur chaque brin pour une antenne de taille 50cm, R_1 représente la plus proche de la source d'alimentation, R_{25} correspond à la résistance la plus éloignée.	153
5.2	Paramètres de simulation pour la simulation GPR.	154
5.3	Paramètres diélectriques du type de sol utilisé dans la simulation.	156
5.4	Propriétés du milieu fractal	156
5.5	Les classes de diamètres pour la simulation des tuyaux.	156
5.6	Exemple d'encodage des étiquettes	169
5.7	Rapport de classification des classes de diamètre via le modèle ResNet-50176	
5.8	Rapport de classification	180
5.9	Meilleurs hyper-paramètres sélectionnés	185
5.10	Combinaison des hyperparamètres optimisés: Classes milieu de propa- gation et type de tuyau.	185
5.11	Combinaison des hyperparamètres optimisés: Classe diamètres.	186
5.12	Rapport de classification sur les données de test- classe diamètre	188
5.13	Rapport de classification classe milieu de propagation	188
5.14	Rapport de classification pour les labels de type de tuyau	189

Introduction Générale

Avec le développement d'internet et des bases de données, on assiste à une évolution exponentielle de la quantité de données disponibles dans tous les domaines et en particulier dans les domaines de la physique. De ce fait, la gestion de la recherche de données dans ces masses d'informations devient difficile et nous avons assisté à l'éclosion de méthodes de fouilles de données (data mining) puis plus récemment l'exploitation de cette mine d'information pour construire des réseaux de neurones spécifiques dits réseaux profonds. Cela a donné naissance à l'Intelligence Artificielle (IA) qui se propose d'aller plus loin quand à la faculté d'apprentissage de ces réseaux par eux même. Nous voyons de plus en plus d'applications de l'IA, notamment la reconnaissance de personnes dans une foule, l'exploitation à des fins commerciales pour faire des propositions de produits adaptés à nos goûts... Si, le domaine de l'IA est en pleine éclosion dans la vie de tous les jours, nous assistons à son utilisation de plus en plus grande dans divers domaines scientifiques. L'IA s'avérant très performante dans le domaine de l'image pour reconnaître des éléments dans une scène, en déduire des similitudes, un des domaines de prédilection en sciences est la médecine où elle peut apporter une aide précieuse au diagnostic. Toutefois, bien d'autres domaines des sciences s'intéressent depuis quelques années à cette évolution sans cesse croissante de l'IA, c'est dans ce contexte que nous projetons dans le cadre de cette thèse de l'appliquer à deux domaines de l'électromagnétisme qui peuvent à priori s'avérer décorrélés mais qui sommes toutes reposent sur un même principe : la réflectométrie dans le domaine des radiofréquences. Ainsi, nous allons voir l'applicabilité des réseaux de neurones profonds à :

- La réflectométrie temporelle monodimensionnelle appliquée au diagnostic de défauts localisées le long de réseaux de câbles.
- La réflectométrie temporelle tridimensionnelle appliquée à la recherche de tuyaux enfouis dans le sous-sol et à leur caractérisation. Il s'agit du domaine du radar à pénétration de sol classiquement appelé GPR pour Ground Penetrating Radar.

Le choix de ces domaines d'application est lié à un besoin sans cesse croissant de techniques performantes pour répondre à des contraintes de plus en plus sévères comme nous allons le détailler par la suite. Considérons tout d'abord la problématique de détection de défauts le long de réseaux complexes de câbles Dans presque tous les domaines industriels, les câbles électriques agissent comme les fondements, en étant les conduits principaux pour acheminer l'énergie et l'information. Leur rôle est incontournable pour assurer le bon fonctionnement de divers équipements et structures,

tels que les automobiles, les machines industrielles, les centrales électriques, et de nombreuses autres infrastructures. Avec l'innovation et l'évolution technologique, ces câbles ont subi d'importantes modifications et améliorations, devenant de plus en plus longs et élaborés pour répondre aux exigences croissantes de diverses applications. Ils sont désormais caractérisés par des conceptions plus sophistiquées, impliquant une complexité accrue dans leur configuration et leur agencement. Néanmoins, cette complexification n'est pas sans risques. Les dommages ou défaillances de ces câbles, surtout ceux dédiés aux systèmes de sécurité et de contrôle, peuvent avoir des conséquences catastrophiques et fatales. On en trouve des exemples tragiques dans plusieurs accidents aériens où des défaillances de câbles ont été impliquées, démontrant la gravité potentielle des défauts dans ces systèmes cruciaux. Dans ce contexte, la réflectométrie s'est révélée être une méthode diagnostic essentielle, permettant d'identifier les défauts et discontinuités d'impédance dans les câbles électriques. Cette technique de diagnostic s'efforce de repérer les anomalies et les points de défaillance dans les réseaux de câbles, contribuant ainsi à maintenir l'intégrité et la fonctionnalité des systèmes. Cependant, malgré la précision de la réflectométrie, la détection de certains défauts, notamment ceux qui sont moins évidents (défauts non francs), s'avère être un défi. Des défauts souvent invisibles ou ignorés avec les tests habituels nécessitent des méthodes de diagnostic plus poussées et fiables pour être repérés avec certitude. En réponse à ces défis, des approches de réflectométrie avancées ont été mises au point, en domaine temporel (TDR), et elles ont été intégrées avec les technologies d'apprentissage profond, comme les réseaux de neurones, pour optimiser la détection et la localisation des défauts dans des configurations de câbles complexes. Cette intégration a permis d'approfondir l'analyse des défauts et d'améliorer la précision du diagnostic, assurant ainsi une sécurité et une fiabilité accrues des systèmes industriels. L'utilisation de techniques avancées de réflectométrie couplées à l'apprentissage profond est cruciale pour assurer l'intégrité des systèmes câblés dans un contexte industriel. Dans ces environnements, la précision et la fiabilité sont indispensables pour éviter des risques et incidents graves dus à des défaillances de câbles. Concernant le second aspect qui touche au domaine du radar à sondage de sous-sol, les travaux entrent dans le cadre de recherche de méthodes non destructives pour l'analyse de sous-sols et la recherche puis la caractérisation d'objets enfouis. Cette problématique est devenue importante dans les domaines de la géologie et de l'archéologie mais aussi dans le domaine du génie civil. En effet, la recherche de canalisations, leur identification est un enjeu majeur notamment lorsqu'on désire creuser le sous-sol et éviter de les endommager. Aussi, des techniques rapides de mesure en temps réel que sont les radars

de type GPR associées à l'analyse en temps réel par des réseaux de neurones profonds permet d'apporter une réponse à cette problématique. C'est dans ce contexte que nous proposons de voir la faisabilité de détection de tuyaux enfouis par Deep Learning. Dans ce travail, une exploration approfondie et une optimisation des réseaux de neurones profonds sont menées, faisant écho aux besoins croissants dans les domaines de la réflectométrie 1D pour le diagnostic des câbles et de la réflectométrie 3D pour l'analyse géophysique via le radar à pénétration de sol (GPR). Ces technologies ont révolutionné les méthodes actuelles, permettant une détection et une classification avancées des défauts dans les câbles et des objets souterrains, en apportant des solutions inédites et efficaces aux problématiques préexistantes. Les objectifs de cette recherche poursuivent deux buts, cherchant à établir un diagnostic précis et fiable des anomalies présentes dans les câbles et à effectuer une classification avancée des données issues du GPR. En focalisant sur l'amélioration de la détection des irrégularités moins apparentes dans les structures câblées complexes et sur l'interprétation des données GPR pour une identification exacte des objets enterrés, ce travail met en lumière le rôle central des réseaux de neurones profonds. Ces derniers se révèlent indispensables pour naviguer à travers les complexités des réflectométries 1D et 3D, promettant une avancée significative dans ces disciplines. Comme vu plus haut, l'objectif principal de cette étude est d'utiliser de nouvelles méthodes basées sur l'intelligence artificielle pour aborder les problèmes liés à l'analyse des défauts dans les câbles et classification des objets enfouis. Ces nouvelles méthodes nous aident à mieux comprendre et gérer les problèmes rencontrés dans les systèmes de câblage et permettent une analyse plus précise des objets enterrés grâce à la technologie GPR. La réalisation de ces objectifs, guidée par les progrès de l'intelligence artificielle, marque un progrès significatif dans les domaines concernés, améliorant la précision et l'efficacité de ces technologies à un niveau supérieur.

Structure de la Thèse

Cette thèse est composée de cinq chapitres. Ils explorent les fondements de l'apprentissage profond, les approches de réflectométrie pour la localisation de défauts, l'application de réseaux de neurones dans le diagnostic de réseaux de câbles, l'utilisation du radar à pénétration de sol et la réflectométrie électromagnétique, et enfin, l'expérimentation et l'adaptation des réseaux de neurones profonds dans ces domaines. Chacun de ces chapitres apporte une contribution à l'amélioration des techniques de diagnostic basées sur la réflectométrie et l'apprentissage profond pour les réseaux de câbles et les radars à pénétration de sol.

- Chapitre 1 : Fondements et Développements de l'Apprentissage Profond Ce chapitre explore les principes fondamentaux et les développements récents des techniques d'apprentissage automatique et profond, se concentrant particulièrement sur les réseaux de neurones profonds qui seront utilisés ultérieurement dans ce travail.
- Chapitre 2 : Approches de Réflectométrie pour la Localisation de Défauts, ce chapitre est consacré à la présentation d'une technique de localisation des défauts basée sur la réflectométrie, mettant l'accent sur les notions de ligne de transmission, les méthodes et les types de signaux injectés, décrivant leur fonctionnement, avantages et limites.
- Chapitre 3 : Application de réseaux de neurones dans le diagnostic de réseaux de Câbles, ce chapitre détaille une méthodologie permettant d'intégrer les réseaux de neurones dans le diagnostic de réseaux de câbles complexes, mettant en avant l'analyse de défauts non-francs et l'utilisation de la réflectométrie dans le domaine temporel.
- Chapitre 4 : Radar à Pénétration de Sol et réflectométrie électromagnétique, ce chapitre explore l'utilisation du radar à pénétration de sol (GPR), décrivant les phénomènes électromagnétiques liés à la propagation des ondes, le principe de fonctionnement du radar GPR, ses composants et son dimensionnement, et introduit la méthode FDTD pour la construction de la base de données.
- Chapitre 5 : Expérimentation et adaptation des réseaux de neurones profonds, ce chapitre final explore l'application et l'adaptation des réseaux de neurones profonds dans le domaine des radars à sondage de sol, détaillant tout le processus, depuis la génération de la base de données jusqu'à la mise en œuvre et l'analyse des performances des différentes architectures de réseaux de neurones développées.

1

Introduction a l'apprentissage automatique et état de l'art sur l'apprentissage profond

Sommaire

1.1	Introduction	23
1.2	Apprentissage automatique	24
1.3	Introduction aux Réseaux de neurones artificiels	34
1.4	Apprentissage des réseaux de neurones artificiels	38
1.5	Réseau de neurones convolutifs	47
1.6	Architectures avancées	60
1.7	Conclusion	63

Ce chapitre présente les principes fondamentaux et les développements récents des techniques d'apprentissage automatique et profond. Nous nous concentrons principalement sur les réseaux de neurones artificiels profonds, qui seront ultérieurement utilisés dans le cadre de ce travail de thèse.

1.1 Introduction

L'apprentissage automatique a donné naissance à une science en pleine expansion: l'intelligence artificielle (IA). Marvin Lee Minsky [1] est l'un des principaux pionniers dans le domaine, il en a donné une définition qui est la suivante: il s'agit de la création de programmes informatiques capables d'effectuer des tâches, communément appelée "construction de programmes informatiques engageant des tâches". De nos jours, les tâches sont exécutées avec une plus grande efficacité par les humains car elles nécessitent des capacités cognitives d'un niveau avancé, y compris le processus d'acquisition de connaissances (processus d'apprentissage). Le trio des capacités cognitives comprend la perception, l'organisation de la mémoire et le raisonnement critique.

La création des principes fondateurs de l'IA est enracinée dans les premières réflexions de quelques penseurs visionnaires. Dans son ouvrage "Computing Machinery and Intelligence", [2] publié dès 1950, Alan Turing et d'autres chercheurs [3] se sont penchés sur le sujet. Turing a exposé les bases de l'intelligence et présenté des contre-arguments aux objections qui contestent sa validité: l'attribut de l'intelligence est exclusif à l'espèce humaine.

L'IA, considérée comme un domaine d'étude fondamental, intègre désormais une gamme d'objectifs. Les disciplines de l'IA, comme l'apprentissage automatique, sont construites sur des idées appelées paradigmes. Deux de ces paradigmes sont les modèles d'apprentissage connexionnistes (qui imitent le cerveau) et les modèles symboliques. Commenant par un aperçu des concepts d'apprentissage automatique, cette thèse se focalise sur l'utilisation des réseaux de neurones artificiels profonds pour des applications dans le domaine de la compatibilité électromagnétique, et du radar. La figure 1.1 illustre les concepts interdépendants de l'IA, de l'apprentissage automatique et de l'apprentissage en profondeur.

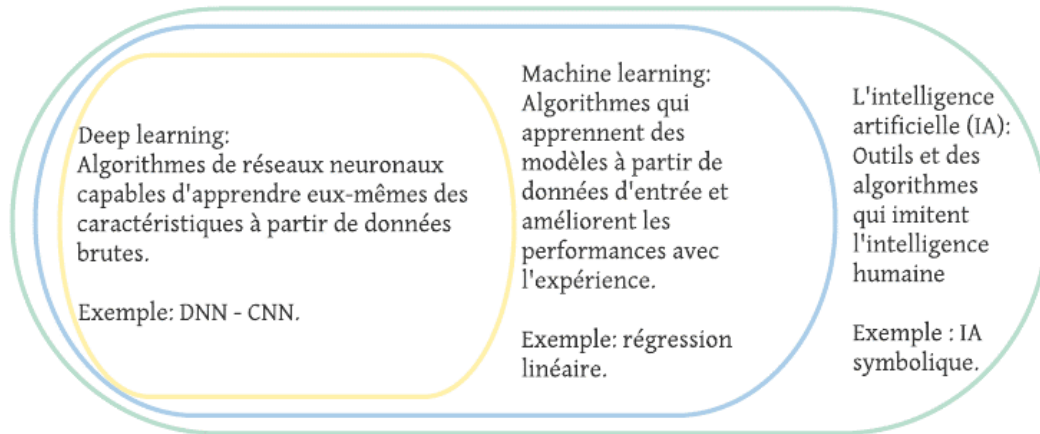


Figure 1.1: Diagramme de Venn représentant les relations des domaines liées à l'IA

1.2 Apprentissage automatique

Un algorithme d'apprentissage automatique est un algorithme qui a la capacité d'apprendre à partir de données. L'américain Tom Mitchell [4] a défini ce qu'est l'apprentissage automatique: une machine apprend lorsque ses performances sur une certaine tâche s'améliorent avec de nouvelles expériences.

De nombreuses définitions ont fait surface dans le domaine de l'apprentissage automatique, notamment:

- Les recherches d'Arthur Samuel [5] en 1959 ont ouvert la voie à un domaine de l'informatique qui offre aux machines la capacité d'apprendre de manière autonome, sans programmation explicite.
- La déclaration de Tom Mitchell de 1998 postule [4] qu'un programme informatique peut apprendre en acquérant de l'expérience E dans une tâche particulière T avec une mesure de performance correspondante P . La capacité du programme à exécuter T , telle que mesurée par P , s'améliore à mesure qu'il rassemble plus de E , conduisant à sa classification en tant que programme d'apprentissage. Si l'expérience E conduit à une amélioration de la mesure de performance P sur la tâche T , alors la performance de la tâche, telle que mesurée par P , est directement liée à E .

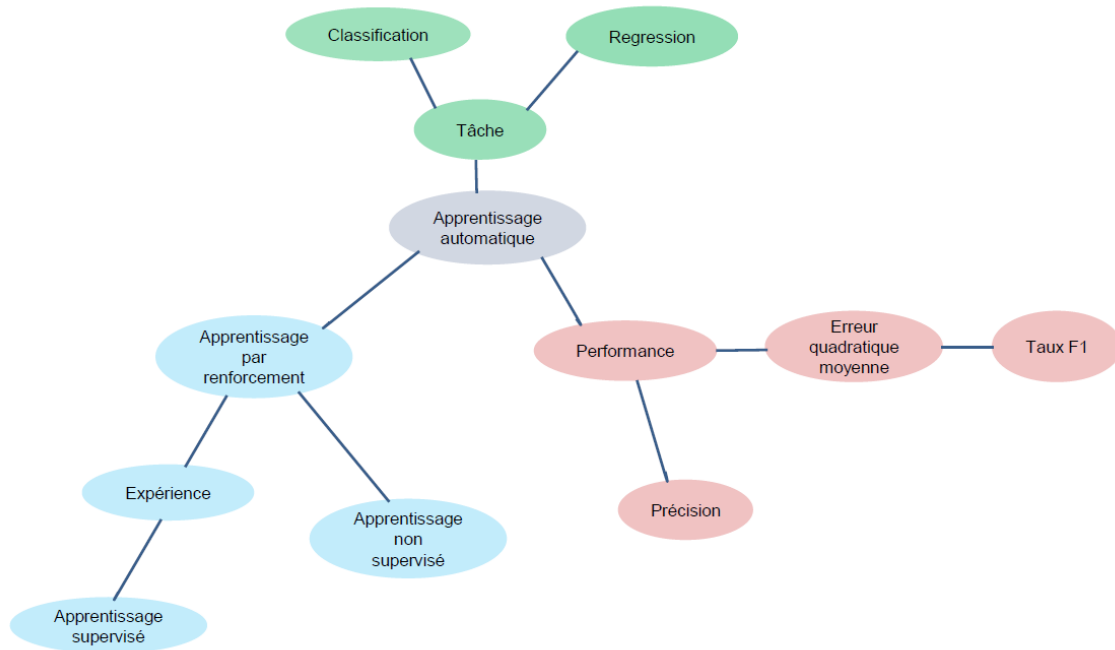


Figure 1.2: Méthodes d'Apprentissage Automatique

Après avoir rappelé ces définitions générales, il devient clair que l'apprentissage automatique [6] peut être décrit comme un sous-domaine de l'intelligence artificielle qui se concentre sur la création, l'analyse et la mise en œuvre de méthodes automatisables permettant aux machines de résoudre des problèmes. Il implique un processus d'acquisition de connaissances et de compétences, permettant ainsi l'exécution de tâches au-delà des capacités initiales de la machine. Au lieu d'être explicitement programmé, l'apprentissage automatique se concentre sur les avancées dans la création de modèles qui permettent une représentation efficace de caractéristiques spécifiques. En tant qu'experts en données, les individus acquièrent une compréhension des caractéristiques statistiques spécifiques des distributions de données. Pour accomplir diverses tâches, le processus utilisé repose sur des modèles construits à partir d'un ensemble d'instances ou d'observations acquises lors de l'apprentissage. La définition de Tom Mitchell [4] sur les mécanismes de bonne généralisation met en évidence un lien solide entre trois concepts clés dans le domaine de l'apprentissage.

1.2.1 Expérience E et tâche T

Il existe deux grandes catégories d'algorithmes d'apprentissage automatique:

- Algorithme supervisé

- Algorithme non supervisé

1.2.2 Apprentissage supervisé

Dans l'apprentissage supervisé [7], chaque échantillon présent dans l'ensemble de données est généralement accompagné d'une bonne réponse prédéterminée et connue. Pour entraîner un modèle d'apprentissage automatique, il est nécessaire de fournir un ensemble de données, composé de points d'observation et leur réponse réelle correspondante. Ces points peuvent provenir aussi bien de simulations numériques que d'expérimentations.

L'apprentissage supervisé permet de répondre à deux objectifs potentiels: la classification dont le but est d'identifier et de reconnaître des classes d'objets par exemple dans une image et la régression [8] qui vise plutôt à estimer des variations d'une fonction continue comme nous allons le détailler par la suite.

1.2.3 Classification

La résolution de problèmes de classification [6] permet de produire des étiquettes, il s'agit d'une valeur discrète attribuée à un objet particulier (exemple chien, chat, auto... sur une image). Binaire et multi-classes sont les deux catégories de la classification, le binaire ayant une sortie 0 ou 1. Plusieurs notations peuvent être employées pour distinguer les classes négatives et positives, mais elles peuvent toujours être ramenées à une sortie binaire (zéro et un). Un ensemble de nombres est utilisé pour exprimer la sortie d'un classificateur multi-classes.

1.2.3.1 Classification binaire

Dans l'apprentissage automatique, la classification binaire consiste à catégoriser les données en deux groupes ou catégories distincts. Les deux groupes peuvent être désignés par positif et négatif, vrai et faux, ou simplement 0 et 1.

$$y = \begin{cases} 1, & \text{si } w \cdot x + b > 0 \\ 0, & \text{sinon} \end{cases} \quad (1.1)$$

Où :

- y est la sortie de la classification (0 ou 1).
- w est le vecteur de poids.
- x est le vecteur d'entrée (caractéristiques de l'objet).
- b est le biais.

Pour ce faire, des modèles de classification binaires [6] sont entraînés à l'aide d'un ensemble de données étiquetées. Ainsi, dans le fichier d'apprentissage, les sorties sont représentées par des 0 ou 1. Le modèle apprend des données pour réaliser des prédictions de l'étiquette permettant ensuite de prédire la catégorie auxquels appartiennent les nouveaux points de données non présents dans la base. Cette technique est souvent utilisée dans diverses applications telles que le diagnostic médical, la détection de spams et la détection de fraude. Par exemple, un modèle de classification binaire pourrait être formé pour trier les e-mails en catégories de spam ou non en reconnaissant des modèles et des mots clés spécifiques. Une fois formé, le modèle peut classer efficacement les nouveaux e-mails comme spam ou non spam.

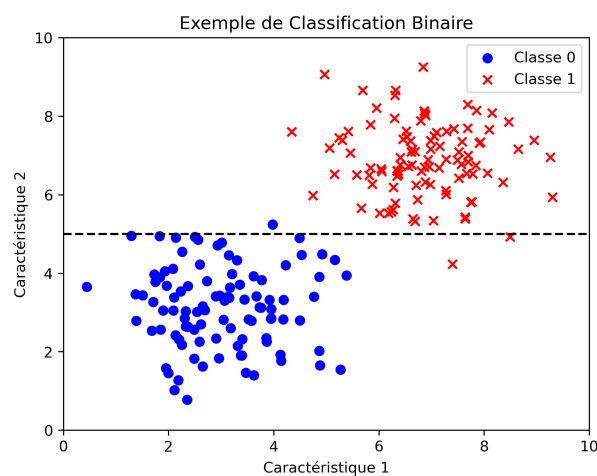


Figure 1.3: Classification binaire

1.2.3.2 Classification multi-classe

L'apprentissage automatique supervisé [6] consiste à classer les données en plusieurs catégories contrairement au cas binaire, de ce fait le nombre de classes peut être supérieur à 2. C'est ce qu'on appelle la classification multi-classes. Dans ce cas, le modèle est destiné à prédire une catégorie et une seule parmi plusieurs options.

Une forme courante est la classification d'images, dans lesquelles différentes classes peuvent catégoriser leur contenu. Pour illustrer, les émotions humaines telles que la joie, le choc, la surprise, ou la colère peuvent être distinguées sur une image. La reconnaissance faciale est également une forme de classification d'images multi-classes. Un autre exemple de classification multi-classes, consiste à diagnostiquer l'état de santé d'un patient et à évaluer un niveau de gravité. L'un des cas d'utilisation les plus courants consiste à identifier la gravité de la rétinopathie en analysant les scanners oculaires du patient [9]. Un système d'apprentissage automatique spécialisé dans la classification multi-classes est utilisé pour déterminer si le patient est atteint de la maladie et si oui, dans quelle mesure, le système classe la gravité de la maladie comme légère, modérée, sévère, ou proliférative en fonction de son analyse. Soit N le nombre total de classes dans la classification multi-classes. Chaque instance de données x est associée à un vecteur \mathbf{y} de longueur N , où chaque composante y_i indique la probabilité que l'instance appartienne à la classe i .

$$\mathbf{y} = [y_1, y_2, \dots, y_N]$$

La prédiction de la classe pour une instance donnée est généralement effectuée en sélectionnant la classe j qui possède la plus grande probabilité y_j :

$$\text{Classe prédite} = \arg \max_j y_j$$

Classification Multi-étiquettes (Multi-label)

En ce qui concerne la classification multi-étiquettes, chaque élément peut appartenir à plusieurs classes. Ceci diffère de la classification typique à étiquette unique qui a été décrite plus haut où les objets ne peuvent appartenir qu'à une seule catégorie. Cette problématique multi-label [10] devient incontournable du fait de la définition d'un nombre de classes importants recoupant différentes propriétés d'un même objet. Outre cette augmentation croissante, les scènes naturelles peuvent être complexes à classer, car une seule image peut appartenir à plusieurs catégories sémantiques. Par exemple, une image peut appartenir simultanément à des catégories telles que la mer, le vent, le soleil, la pluie, etc. Des exemples de classes sémantiques incluent des phénomènes naturels comme les océans et les couchers de soleil. L'équation illustrative pour le cas de la classification multi-étiquettes pourrait être formulée comme suit:

$$Y = f(X)$$

où:

- Y représente l'ensemble des étiquettes ou classes auxquelles un élément peut appartenir. Dans le contexte de la classification multi-étiquettes, Y serait un ensemble de variables binaires indiquant la présence ou l'absence de chaque classe.
- X représente les caractéristiques ou attributs de l'élément à classifier.
- f est une fonction de classification qui attribue les étiquettes Y en fonction des caractéristiques X .

Dans ce cas, Y est un vecteur de taille N , où N est le nombre total de classes possibles. Chaque composante Y_i du vecteur indique si l'élément appartient à la classe i (1) ou non (0).

1.2.4 Régression

La régression [6] [11] est une approche statistique permettant de déterminer la relation entre des variables dépendantes en général continues et des variables indépendantes. En termes plus simples, il évalue le lien entre des variables « réponses » avec une ou plusieurs variables caractéristiques d'entrée. Il apparaît que la problématique est différente du cas précédemment décrit puisque les données à traiter (ainsi que les résultats) sont continues et non plus discrètes. De nombreuses techniques permettent d'effectuer ce type d'opération. La méthode d'analyse de régression largement utilisée est la régression linéaire. La régression linéaire [11] est une technique d'apprentissage automatique largement utilisée qui cherche à établir une corrélation linéaire entre une ou plusieurs variables indépendantes (utilisées pour prédire la variable dépendante) et la variable dépendante elle-même (la variable à prédire). Il est particulièrement utile dans les scénarios prédictifs, tels que l'estimation de la valeur d'une maison, où l'objectif est de prédire une valeur numérique continue basée sur diverses variables d'entrée. La régression linéaire sert d'algorithme fondamental [12] pour des modèles de prédiction plus sophistiqués dans l'apprentissage automatique. Sa facilité de mise en œuvre et d'interprétation, associée à son efficacité pour divers problèmes de prédiction, en font un choix classique.

Le principe de base de la régression linéaire pour une seule variable d'entrée est d'identifier une droite qui minimise l'écart entre les points de données réels et ceux de la droite de régression. Cette droite est caractérisée par sa pente et son ordonnée à l'origine. L'objectif principal de la régression linéaire est de déterminer les valeurs de ces coefficients qui minimisent l'erreur de prédiction.

Pour illustrer nos propos, nous allons brièvement décrire le principe de la régression

linéaire simple qui consiste à identifier les paramètres d'une droite approximant au mieux un nuage de points dans un plan xy lorsque l'on ne considère qu'une variable d'entrée et une variable de sortie. L'équation ci-dessous représente la droite de régression:

$$y = \beta_0 + \beta_1 x + \epsilon \quad (1.2)$$

où :

- y est la variable dépendante (ce que nous essayons de prédire),
- x est la variable indépendante,
- β_0 est l'ordonnée à l'origine,
- β_1 est la pente de la droite,
- ϵ est l'erreur (la différence entre la valeur réelle et la valeur prédite).

Le processus de minimisation vise à réduire l'erreur entre les valeurs prédites et les valeurs réelles. Cette erreur est appelée "résidu", et est notée " ϵ ". Dans le cas où nous avons un vecteur d'entrée de variables indépendantes, l'objectif est de trouver les paramètres de l'hyperplan qui minimisent cette erreur.

1.2.5 Apprentissage non supervisé

L'apprentissage non supervisé [13] est un processus d'apprentissage automatique qui fonctionne sur un ensemble de données non étiquetées, permettant aux modèles d'agir dessus de manière indépendante. Contrairement à l'apprentissage supervisé, l'apprentissage non supervisé n'a pas de données de sortie correspondantes pour les données d'entrée dans les problèmes de régression ou de classification. Son objectif est d'identifier la structure intrinsèque de l'ensemble de données, de regrouper les éléments de données similaires et de les représenter sous une forme condensée. Les problèmes d'apprentissage non supervisé peuvent être classés en deux catégories principales suivant: les règles de regroupement et les règles d'association.

1.2.5.1 Regroupement

La technique de regroupement consiste à organiser des objets en groupes en fonction de leurs similitudes, en veillant à ce que les éléments similaires soient regroupés, tandis que ceux qui ne partagent pas de similitudes sont séparés. Afin de classer les objets de données, l'analyse de cluster identifie les points communs entre eux et les regroupe en fonction de la présence ou de l'absence de ces similitudes.

1.2.5.2 Association

Les règles d'association établissent des corrélations entre les variables au sein de bases de données étendues. Ces règles identifient la combinaison spécifique d'éléments qui apparaissent souvent ensembles dans les données. Par exemple, l'application de règles d'associations à une campagne marketing améliore son efficacité globale. Ainsi, si un client achète du pain (élément X), il peut également être plus susceptible d'acheter du beurre ou de la confiture (élément Y). L'analyse du panier de consommation est l'une des applications les plus courantes des règles d'association.

1.2.6 Performance P

Pour évaluer l'efficacité d'un modèle d'apprentissage automatique de nombreuses méthodes peuvent être employées. Cette partie donne un aperçu des principales métriques utilisées dans des tâches de classification et de régression [14] [15].

1.2.6.1 Classification

Définition de matrice de confusion et métrique statiques

L'évaluation de la qualité d'un modèle prédictif implique généralement l'examen de ses erreurs de classification. Cependant, il est important de noter qu'il est souvent préférable d'exprimer les erreurs en proportion du nombre total d'observations. La matrice de confusion permet de mesurer la capacité du modèle à prédire correctement les classes des échantillons.

La matrice de confusion est une matrice qui affiche les nombre d'observation qui ont été bien classifiées et incorrectement classé par le modèle. Dans le cas de classification binaire, la matrice de confusion est représentée comme suit:

		Classe réelle	
		-	+
Classe prédite	-	Vrais négatifs	Faux négatifs
	+	Faux positifs	Vrais positifs

Figure 1.4: Classification binaire

- Les vrais positifs (**TP** : True Positive en anglais) représentent les échantillons correctement classés comme positifs.
- Les faux négatifs (**FN** : False Negative) représentent les échantillons incorrectement classés comme négatifs.
- Les faux positifs (**FP** : False Positive) représentent les échantillons incorrectement classés comme positifs.
- Les vrais négatifs (**TN** : True Negative) représentent les échantillons correctement classés comme négatifs.

En utilisant une matrice de confusion, on peut calculer diverses métriques d'évaluations d'un modèle de classification, notamment la précision, le rappel, et le score F_1 .

Précision:

C'est la mesure la plus simple et la plus courante pour évaluer la performance d'un algorithme de classification. La précision explique combien de cas prédits correctement se sont révélés positifs.

La précision d'une étiquette est définie comme *le nombre de vrais positifs divisé par le nombre de positifs prédits*.

$$\text{Précision} = \frac{TP}{TP + FP} \quad (1.3)$$

L'exactitude (Accuracy):

La précision mesure simplement la fréquence des prédictions correctes du classificateur. Nous pouvons définir la précision comme le rapport entre le nombre de prédictions correctes et le nombre total de prédictions.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (1.4)$$

Rappel (Recall):

Cette métrique représente le nombre de vrais positifs sur le nombre total de vrais positifs et faux négatifs. Elle mesure le pourcentage d'exemples positifs que le modèle a correctement identifiés par rapport au nombre total d'exemples positifs dans l'ensemble de données.

Un rappel élevé indique que le modèle a une bonne capacité à détecter les exemples positifs, tandis qu'un faible rappel indique que le modèle n'a pas la capacité à détecter

les échantillons bien classifiés. Le rappel d'une étiquette est défini *comme le nombre de vrais positifs divisé par le nombre total de vrais positifs et de faux négatifs*.

$$Rappel = \frac{TP}{TP + FN} \quad (1.5)$$

F_1 _mesure (F_1 _score):

C'est une métrique de performance qui représente la moyenne harmonique de la précision et le rappel:

$$F_1_score = 2 \frac{Precision \cdot Recall}{Precision + Recall} \quad (1.6)$$

Les valeurs du score F_1 se situent sur une échelle de 0 à 1. Une valeur proche de 1 représente une performance supérieure tandis qu'une valeur proche de 0 indique une performance inférieure. Dans les tâches de classification où les proportions de classe sont déséquilibrées, le score F_1 apparaît comme une mesure plus précise que la précision. Ceci est illustré lorsque la classe positive représente un pourcentage minimal de l'ensemble de données total, ce qui peut entraîner un score de précision plus élevé malgré le fait que le modèle ne détecte que quelques exemples positifs. Le score F_1 tient compte à la fois de la précision et du rappel, fournissant une évaluation complète des performances globales du modèle dans de tels cas.

1.2.6.2 Régression

Dans le cas d'un problème de régression [16], et avant d'aborder les principales mesures d'évaluations, il est nécessaire de définir ce que signifie un "résiduel" lors d'une tâche de régression. Le résidu peut être défini comme la différence entre les valeurs réelles et les valeurs prédites.

Erreur absolue moyenne (MAE):

L'erreur moyenne absolue (MAE) représente la différence entre les valeurs originales et les valeurs prédites extraites en calculant la moyenne de la différence absolue sur l'ensemble des données. Ainsi pour N points, si y_i représente la valeur vraie et y_i_tilde l'estimation, la MAE est donnée par: L'erreur moyenne absolue (MAE) représente la différence entre les valeurs originales et les valeurs prédites, obtenue en calculant la moyenne de la différence absolue sur l'ensemble des données. Ainsi, pour N points, si

y_i représente la valeur vraie et \hat{y}_i l'estimation, la MAE est donnée par :

$$MAE = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i| \quad (1.7)$$

Erreur quadratique moyenne (MSE):

L'erreur quadratique moyenne (MSE) représente la moyenne des différences entre les valeurs réelles et les valeurs prédites au carré.

$$MSE = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2 \quad (1.8)$$

Racine Carrée de l'Erreur Quadratique Moyenne (RMSE):

Il s'agit de la racine carrée de l'erreur quadratique moyenne

$$RMSE = \sqrt{MSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2} \quad (1.9)$$

1.3 Introduction aux Réseaux de neurones artificiels

1.3.1 Retour sur les réseaux de neurones humains

Le cerveau humain est dominé par deux types de cellule : les neurones et les cellules gliales. Les neurones, au nombre de 86 à 100 milliards, sont supportés par des cellules gliales qui facilitent leur fonctionnement et assurent leur protection. Chaque neurone peut former de 1 000 à 10 000 connexions synaptiques. La structure neuronale est constituée d'un soma central, d'un noyau, d'axones et de dendrites. Ces axones peuvent s'étendre sur plusieurs mètres, permettant la transmission de signaux à longue distance. Les synapses sont les points de connexion entre neurones, facilitant la communication. Les neurones biologiques intègrent une multitude de stimuli et déclenchent des potentiels d'action en réponse. Les réseaux neuronaux artificiels en apprentissage automatique imitent la structure et la fonctionnalité des neurones biologiques . Ils intègrent des entrées multiples (similaires aux stimuli pour les neurones biologiques) et génèrent une sortie après traitement, analogue au potentiel d'action

dans le contexte biologique. Un exemple illustrant l'analogie entre le neurone biologique et le neurone artificiel est présenté à la Figure 1.5.

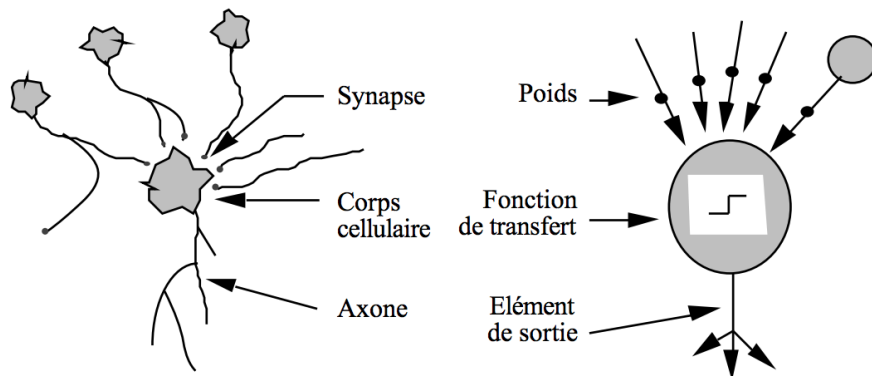


Figure 1.5: Analogie entre le neurone biologique et neurone artificiel [17]

1.3.2 Neurones artificiels (neurone formel)

Le neurone artificiel est un modèle largement inspiré par les neurones biologiques. Sa conception s'est affinée en intégrant des concepts mathématiques et en s'inspirant d'autres domaines d'études. Dans le contexte biologique, le terme "activation" décrit le moment où un neurone est influencé par d'autres neurones conformément à des principes donnés. Quand un neurone atteint son seuil d'activation, il transmet un signal aux autres neurones. Dans les réseaux neuronaux, cette connexion est symbolisée par une fonction d'activation, laquelle, avec la connexion synaptique, reflète l'influence du neurone émetteur sur le neurone récepteur. La force de cette influence est déterminée par son impact sur la capacité d'apprentissage, et est directement associée à l'ajustement des poids synaptiques. Ces principes fondamentaux sont au cœur de la modélisation neuro-mimétique.

Le modèle de neurone biologique simplifié est capable de calculs rapides. Ceci est illustré à la Figure 1.6, où le neurone reçoit des entrées x . Les poids synaptiques w_j jouent un rôle crucial dans la mise en forme des connexions, conduisant finalement au calcul d'une somme par l'additionneur. Les entrées du vecteur d'entrée x et du vecteur de poids w sont combinées et traitées pour produire la sortie, notée y . Cela implique de pondérer les entrées de manière appropriée avant de calculer la valeur de sortie. Si cela est jugé nécessaire, un biais peut être ajouté avant que la somme pondérée soit passée à une fonction. Le calcul de la sortie d'un neurone est déterminé par sa fonction d'activation. Une telle fonction revêt une importance considérable.

Chaque paramètre est souvent exprimé dans un vecteur où:

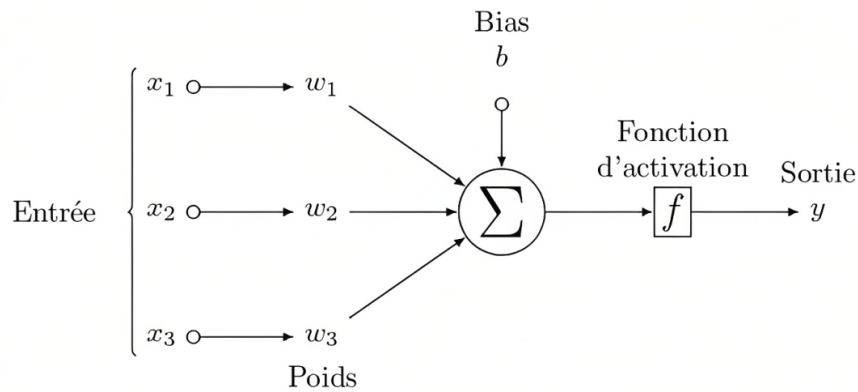


Figure 1.6: Fonctionnement d'un Neurone Artificiel

- $x = (x_0, \dots, x_m)$ représente le vecteur d'entrée
- $w = (w_{k_1}, w_{k_2}, \dots, w_{k_m})$ représente le vecteur de poids associé a chaque entrée
- b_k est le biais
- $f(.)$ est la fonction d'activation

$$f(x) = b + \sum_i w_i \cdot x_i$$

- $y_k = (y_1, \dots, y_s)$ est le vecteur signal de sortie de neurone déterminé comme suit:

$$y = f\left(\sum_{i=1}^n w_i \cdot x_i + b\right)$$

La fonction d'activation $f()$ est une fonction non linéaire permettant d'introduire de la non-linéarité dans un réseau de neurones artificiels. Elle est essentielle pour que les neurones puissent modéliser des relations complexes entre les entrées et les sorties. Les fonctions les plus utilisées seront décrites plus loin dans ce mémoire

Un assemblage de neurones synthétiques est appelé Réseau de Neurones Artificiels (ANN). Un réseau de neurones artificiels (ANN) est composé de neurones connectés à d'autres neurones et régis par un ensemble de règles définies par son architecture et sa destination. Pour former les ANN, des stratégies d'apprentissage sont mises en œuvre telles que calculer les poids synaptiques. Il existe de nombreuses architectures de réseaux. La classification de L'ANN est basée sur des facteurs distinctifs qui incluent:

- Le type de neurone qui peut être binaire, pulsé, cellulaire ou autre.
- Types de topologie de réseau: en couches, récurrent et autres.
- Mécanismes utilisés en fonctionnement: propagation vers l'avant, convolution et autres processus connexes.

1.3.3 Historique et évolution des réseaux de neurones

Origines et débuts:

- En 1943, Warren S. McCulloch et Walter Pitts [18] ont introduit le concept du neurone de McCulloch-Pitts. Bien que ce modèle ne possédait pas de mécanismes d'apprentissage formels, il établissait les bases du paradigme connexionniste de l'IA.
- En 1949, Donald Hebb [19] a introduit la "règle de Hebb" qui indique que si deux neurones s'activent simultanément de manière répétée, leur connexion se renforce. Cette règle est essentielle à l'apprentissage dans les réseaux neuronaux, où les connexions entre neurones sont ajustées en fonction de leur activité.

Développements majeurs:

- À la fin des années 1950, Frank Rosenblatt [20] a introduit le Perceptron, le premier modèle de neurone artificiel pouvant apprendre de l'expérience, inspiré du système visuel humain.
- cependant, en 1969, Minsky et Papert [21] ont souligné certaines limitations du Perceptron, en particulier son incapacité à traiter des problèmes non linéaires, tels que la fonction XOR (OU exclusif).

Renaissance des ANN:

- Après une période de déclin au début des années 1970, l'intérêt pour les ANN a été ravivé. En 1974, Paul Werbos [22] a présenté la rétropropagation, qui a renouvelé l'intérêt pour les ANN, en particulier les réseaux multicouches, connus sous le nom de Perceptrons multicouches (MLP).

Innovations modernes:

- Les années 1990 ont vu l'émergence des réseaux convolutifs (CNN) avec Yann LeCun [23] parmi les pionniers. Il a développé "LeNet" en 1998 pour la classification des caractères manuscrits.
- Par ailleurs, en 1997, Sepp Hochreiter et Jürgen Schmidhuber ont introduit l'architecture LSTM (Long Short-Term Memory), qui a apporté des améliorations significatives dans le traitement des séquences [24].

1.4 Apprentissage des réseaux de neurones artificiels

1.4.1 Définition d'un perceptron multicouche (MLP)

Le perceptron multicouche (MLP) [25] constitue un type répandu de réseau de neurones artificiels, comme illustré à la Figure 1.7. Le réseau comprend plusieurs couches qui facilitent le flux de données uniquement de la couche d'entrée vers la couche de sortie. Il s'agit essentiellement d'un réseau de propagation directe à anticipation composé d'un nombre variable de neurones dans chaque couche. Le nombre de neurones dans la couche d'entrée étant fixé par le nombre d'entrée, de même pour la couche finale. Cette dernière couche, appelée « sortie », est constituée de neurones qui fonctionnent comme les sorties du système global. Un réseau de neurones comprend trois types de couches : une couche d'entrée, qui reçoit les données initiales ; des couches cachées, qui traitent les données reçues d'une couche précédente et les transmettent à la couche suivante ; et enfin une couche de sortie, qui génère une prédiction en fonction du type de problème abordé

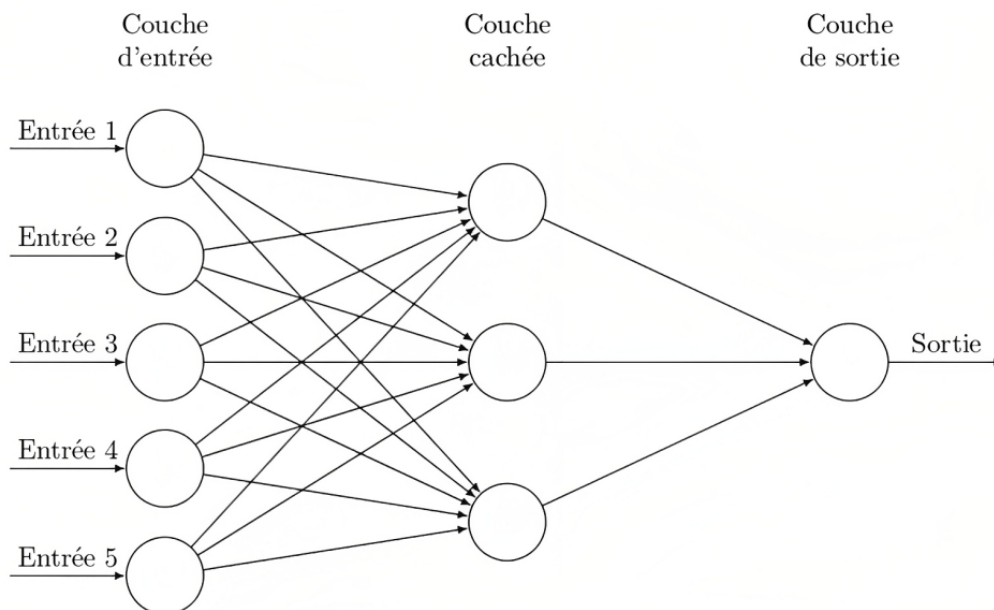


Figure 1.7: Perceptron multicouche (MLP)

Propagation des données

Chaque couche l du MLP transforme l'entrée a^{l-1} de cette couche via une combinaison linéaire suivie d'une activation non-linéaire:

$$z_j^l = \sum_i w_{ji}^l a_i^{l-1} + b_j^l \quad (1.10)$$

$$a_j^l = f(z_j^l) \quad (1.11)$$

où :

- w_{ji}^l est le poids entre le neurone i de la couche $l-1$ et le neurone j de la couche l ,
- b_j^l est le biais du neurone j de la couche l ,
- f est une fonction d'activation (par exemple, sigmoïde, ReLU, etc.)

Couches du MLP

$$a_i^0 = x_i \quad (\text{pour la couche d'entrée}) \quad (1.12)$$

$$a_j^l = f\left(\sum_i w_{ji}^l a_i^{l-1} + b_j^l\right) \quad (\text{pour les couches cachées}) \quad (1.13)$$

$$y_k = a_k^L \quad (\text{pour la couche de sortie}) \quad (1.14)$$

où :

- x_i est la valeur d'entrée i ,
- L est le nombre total de couches dans le MLP,
- y_k est la prédiction de sortie pour le neurone k .

1.4.2 Algorithme de descente de gradient

L'apprentissage consiste à aligner les poids du réseau $w_{i,j}$ avec les données présentées. Afin d'affiner les réponses reçues, des ajustements sont effectués en fonction des apports de la phase d'apprentissage. Le processus de minimisation de la fonction de perte consiste à déterminer les valeurs optimales pour les poids $w_{i,j}$.

Ce processus implique un mécanisme en deux phases:

- **Phase de propagation vers l'avant (feed-forward propagation):** le réseau reçoit les entrées et les transmet à travers les couches neuronales jusqu'à la sortie. Les résultats sont ensuite comparés aux sorties attendues et l'erreur correspondante est calculée
- **Phase de rétropropagation:** [26] afin de bien ajuster les poids des neurones, l'erreur est propagée à travers le réseau neuronal. Ce processus est répété

jusqu'à ce que l'erreur devienne minimale. Dans ce cas, deux types de signaux peuvent être identifiés à travers cette figure 1.7: signaux de fonction; un signal de fonction est un signal d'entrée (stimulus) qui arrive à l'entrée du réseau et se propage vers l'avant (neurone par neurone) à travers le réseau, et sort à l'extrémité du réseau comme un signal de sortie et signal d'erreur. Les signaux d'erreur sont générés par un neurone de sortie du réseau et se propagent en arrière couche par couche à travers tout le réseau.

Descente de gradient

Le processus d'apprentissage [27] d'un réseau de neurones vise à adapter ses poids pour obtenir les meilleures performances, que ce soit en régression ou en classification. L'objectif principal est de minimiser la fonction de coût, ψ , qui mesure l'écart entre les données réelles, notées y , et les prédictions du réseau, \hat{y} .

$$\psi(p) = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2 \quad (1.15)$$

Pour la régression, la fonction de coût couramment utilisée est l'erreur quadratique moyenne (MSE), tandis que pour la classification, l'entropie croisée est fréquemment employée.

L'optimisation vise à identifier les poids qui minimisent cette fonction. Généralement, on utilise la descente de gradient qui ajuste les poids en fonction de la pente de la fonction de coût par rapport à ces derniers. Elle se base sur un taux d'apprentissage pour contrôler l'amplitude des mises à jour.

L'algorithme de rétropropagation, qui emploie la technique de descente de gradient, est l'une des méthodes d'optimisation les plus utilisées. Il permet le calcul du gradient du réseau à l'aide d'un autre réseau connecté en sens inverse du réseau initial.

Considérons un réseau de neurones avec $l + 1$ couches. Soit a^k l'activation (ou sortie) de la couche k . En utilisant la fonction définie précédemment, nous notons f^k comme la fonction de transfert (souvent une fonction d'activation comme ReLU, tanh ou sigmoid) appliquée à l'entrée z^k .

L'opération dans chaque couche pendant la propagation avant peut être décrite comme:

$$z^{k+1} = a^k W^{k+1} + b^{k+1} \quad (1.16)$$

$$a^{k+1} = f^{k+1}(z^{k+1}) \quad (1.17)$$

où W^{k+1} et b^{k+1} sont respectivement les poids et les biais associés à la couche $k + 1$. Pendant la phase d'apprentissage, le réseau propage l'entrée a^0 vers la sortie a^l en suivant la récurrence définie précédemment. Après avoir obtenu la sortie, l'erreur est calculée, souvent à l'aide d'une fonction de coût L , que:

$$\text{Erreur} = L(a^l, y) \quad (1.18)$$

où y est la sortie désirée.

La rétropropagation commence alors par le calcul du gradient de cette erreur par rapport à chaque poids du réseau. L'utilisation de la règle de la chaîne permet de calculer le gradient pour chaque couche en fonction de la couche suivante, en commençant par la dernière couche.

Supposons que δ^k représente le gradient de l'erreur par rapport à l'entrée pondérée de la couche k . Pour une couche donnée, ce gradient est donné par:

$$\delta^k = \left(\frac{\partial f^k}{\partial z^k} \right) \times \left((W^{k+1})^T \delta^{k+1} \right) \quad (1.19)$$

où $\frac{\partial f^k}{\partial z^k}$ représente la dérivée de la fonction d'activation.

Avec ces gradients, les poids et les biais peuvent être mis à jour à l'aide d'une méthode d'optimisation, typiquement une variante de la descente de gradient.

1.4.3 Types de fonctions d'activation

La fonction d'activation [28] dans un réseau neuronal est cruciale pour garantir sa performance. Elle détermine l'activation ou la désactivation des neurones en fonction des valeurs d'entrée, et contrôle la capacité d'apprentissage et les performances du réseau. Elle est essentielle pour apprendre des modèles complexes et non linéaires. Sans elle, la sortie serait une simple fonction linéaire, limitant la capacité d'apprentissage. La fonction d'activation permet au réseau de simuler des données complexes et d'améliorer la classification en convertissant l'espace caractéristique. En résumé, elle joue un rôle fondamental en permettant l'apprentissage de modèles complexes, l'amélioration de la capacité de classification et la modélisation de relations non linéaires.

Les fonctions d'activation sont nombreuses et se déclinent en deux grandes catégories:

- les fonctions d'activation qui sont généralement bornées, $f : \mathbb{R} \rightarrow (a, b)$

- les fonctions d'activation non bornées, illustrées à la Figure 1.8:

$$f : \mathbb{R} \rightarrow (a, \infty) \text{ ou } f : \mathbb{R} \rightarrow (-\infty, \infty)$$

- a et b non infini, généralement entre -1, 0 ou 1. La Figure 1.8 présente les fonctions d'activation bornées (Sigmoide, tanh).

Fonctions d'activations utilisées dans les couches cachées

Nous distinguons 6 types de fonctions d'activation utilisés dans les couches cachées

- Fonction d'activation linéaire rectifié (ReLU)
- Leaky ReLU
- Softplus
- Fonction d'activation logistique (Sigmoide)
- Fonction d'activation tangente hyperbolique (Tanh)
- ELU (Exponential Linear Unit)

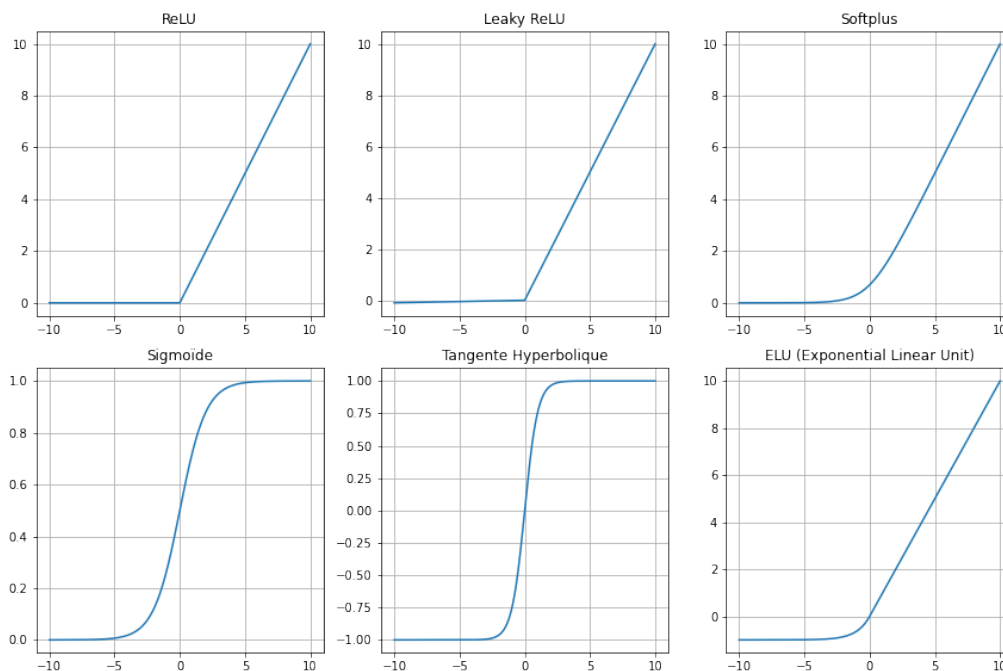


Figure 1.8: Illustration des fonctions d'activation bornées et non bornées.

ReLU (fonction d'activation rectifié): il s'agit d'une fonction non linéaire, qui s'exprime comme suit:

$$f(x) = \max(0, x) \tag{1.20}$$

Cette fonction est importante pour les couches cachées, car elle introduit la non-linéarité dans les modèles ce qui la rend capable d'extraire les caractéristiques les plus pertinentes. La seconde caractéristique de cette fonction est de posséder un gradient non saturant ce qui signifie que le réseau est capable d'apprendre plus rapidement et plus précisément. La fonction d'activation ReLU introduit dans les réseaux neuronaux la sparsité, ce qui implique la suppression des entrées négatives et la réduction du nombre de neurones actifs dans le réseau.

Fonctions d'activations utilisées dans les couches de sorties

Sigmoïde

Il s'agit d'une fonction non-linéaire appelée aussi la fonction logistique car elle est très utilisée dans les modèles de régression logistique. Cette fonction convertit son entrée en une probabilité de valeur comprise entre 0 et 1. Les valeurs négatives élevées sont converties en 0 tandis que les valeurs positives sont converties en 1. Cette fonction a été très utilisée auparavant dans les couches cachées des réseaux multi perceptron et réseaux de neurones convolutifs, mais malheureusement nécessite des calculs lourds.

$$\sigma(x) = \frac{1}{1 + e^{-x}} \quad (1.21)$$

Fonction d'activation tangente hyperbolique

Il s'agit d'une simple extension de la sigmoïde logistique. La seule distinction est que la sortie est comprise entre -1 et 1. Cette fonction permet de mieux faire converger l'algorithme de rétropropagation grâce à sa sortie à large spectre.

Nous allons maintenant nous intéresser au choix du type de fonction d'activation suivant le problème à traiter. En effet, nous avons vu plus haut que les réseaux de neurones étaient à même de résoudre des problèmes de régression et de classification. Le choix pertinent de la fonction doit alors être réalisé. La Figure 1.9 montre les fonctions usuellement utilisées par type de problème.

$$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (1.22)$$

Softmax: Utilisé pour les tâches de classification multiclasse. Il donne une distribution

de probabilité sur plusieurs classes.

$$\text{Softmax}(x)_i = \frac{e^{x_i}}{\sum_j e^{x_j}} \quad (1.23)$$

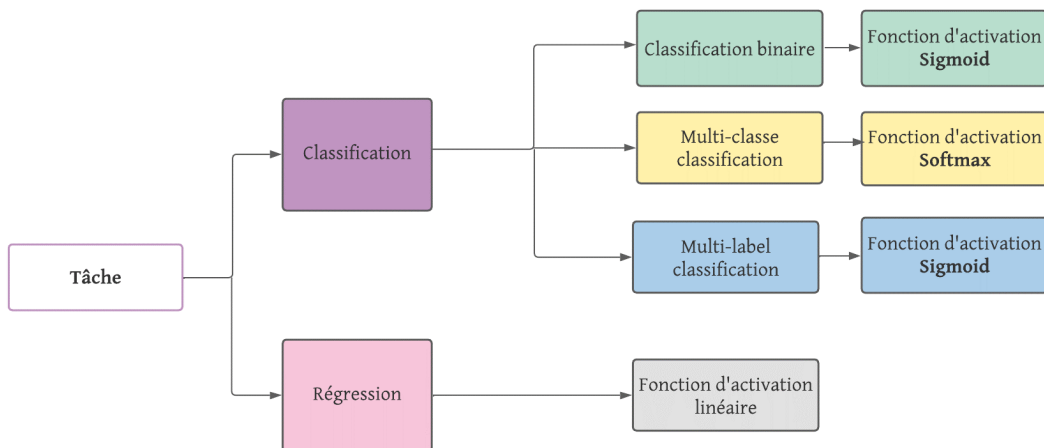


Figure 1.9: Fonctions d'activation selon la tâche : Classification vs Régression

1.4.4 Paramétrage d'un modèle d'apprentissage profond

Cette partie traite une notion essentielle lors de la mise en œuvre d'un modèle d'apprentissage profond: le paramétrage d'un modèle. Bien qu'il n'existe pas de règles fixes pour contrôler et adapter un tel système d'apprentissage profond à une tâche définie, un ensemble d'éléments est à l'ajuster pour obtenir un système performant pour une tâche données et un type d'application définie.

Différence entre paramètres et hyperparamètres

L'optimisation des paramètres est un processus automatique qui permet d'ajuster les poids et les biais du réseau neuronal. Tandis que le paramétrage des hyperparamètres (ou optimisation des hyperparamètres) est un processus manuel qui permet de trouver la meilleure configuration d'hyperparamètres afin de rendre le modèle le plus performant possible. De nombreux hyper-paramètres doivent être définis, par exemple, l'architecture et la taille du réseau, la fonction d'activation utilisée dans chaque couche, le taux d'apprentissage, la fonction de perte, et tous autres paramètres qui permettent d'optimiser les poids de réseau [12] [29].

Capacité, surajustement et sous-ajustement

Le principal enjeu de l'apprentissage automatique est de pouvoir être performant sur de nouveaux éléments, et pas uniquement sur ceux qui ont servi à l'apprentissage de notre modèle. C'est ce qu'on appelle la généralisation; le modèle doit être capable de se généraliser sur des données non vues.

Les modèles d'apprentissage profond peuvent être confrontés à deux principaux problèmes : le surajustement (overfitting) et le sous-ajustement (underfitting), comme illustré à la Figure 1.10. Le surajustement se produit lorsque le modèle est si bien entraîné sur les données d'entraînement qu'il ne parvient pas à généraliser efficacement sur de nouvelles données. Dans ce contexte, plutôt que d'apprendre, le modèle mémorise les données, ce qui limite sa capacité d'apprentissage. À l'inverse, on parle de sous-ajustement lorsque le modèle ne parvient pas à apprendre correctement, résultant en de faibles performances sur l'ensemble d'entraînement. Une des causes de ce problème peut être que le modèle n'est pas assez complexe pour répondre aux besoins spécifiques de la tâche. En augmentant la taille du réseau — le nombre de couches cachées et le nombre de neurones dans chaque couche — les performances du réseau de neurones pourraient être optimisées.

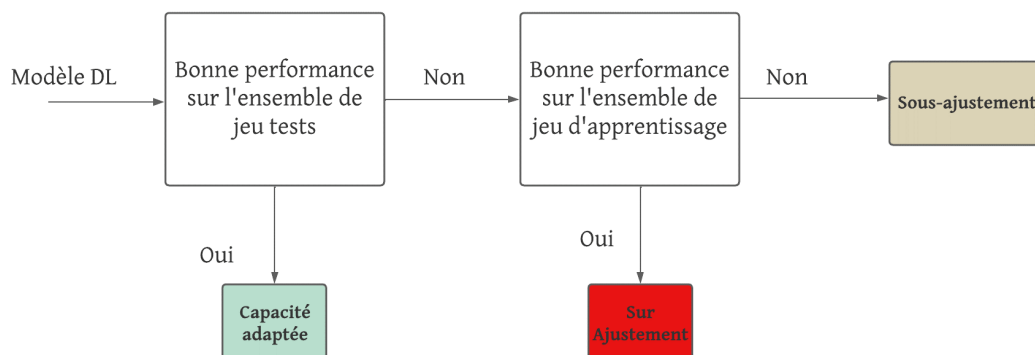


Figure 1.10: Capacité, surajustement et sous-ajustement

Un bon modèle Deep Learning est bien adapté lorsqu'il est performant sur l'ensemble du jeu test. On dit qu'un modèle est dans un régime de surajustement lorsqu'il est performant sur l'ensemble du jeu d'apprentissage uniquement. Un modèle est en sous-ajustement lorsque le modèle n'est pas en mesure d'apprendre sur l'ensemble de jeu d'apprentissage, comme le montre la Figure 1.11.

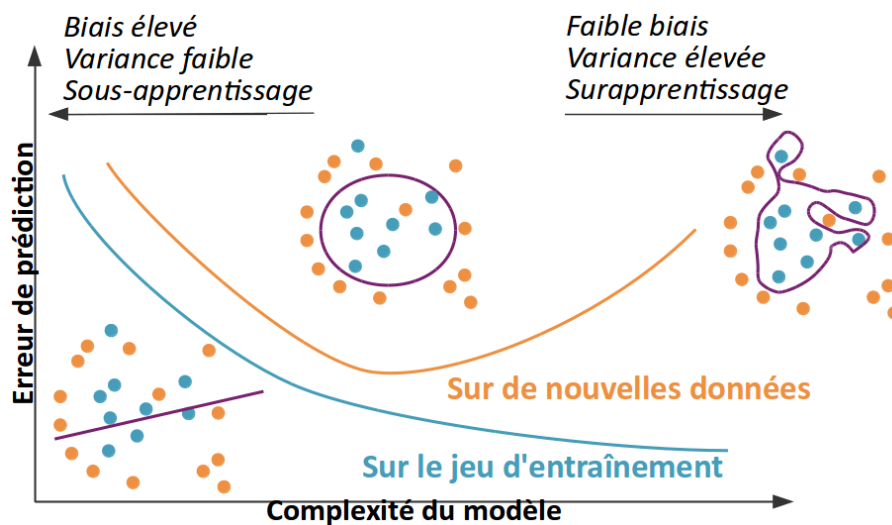


Figure 1.11: Relation typique entre l'erreur de prédiction et la complexité du modèle

Les erreurs de prédictions sur l'ensemble du jeu d'apprentissage et de test se comportent différemment. À l'extrémité gauche du graphique, l'erreur d'apprentissage et l'erreur de généralisation sont toutes les deux élevées. Cette zone est appelée régime de sous apprentissage. Au fur et à mesure que nous augmentons la complexité du modèle, l'erreur d'entraînement diminue mais l'écart entre l'erreur d'apprentissage et l'erreur de généralisation (sur de nouvelles données) augmente. Finalement, l'ampleur de cet écart l'emporte sur la diminution de l'erreur d'apprentissage et nous entrons dans le régime de sur-apprentissage, où la complexité du modèle est trop grande et au-dessus de la complexité optimale.

Pour limiter les problèmes dûs au sur-ajustement, les techniques couramment utilisées sont [6] :

- Augmenter l'ensemble du jeu de données d'entraînement.
- Réduire la taille du réseau et sa complexité.
- Augmenter artificiellement des données (Data Augmentation): [30] cette technique consiste à modifier les données d'entraînement de manière aléatoire (rotation, mise à l'échelle, zoom).
- Utiliser des techniques de la régularisation L_1 et L_2 , le but de ces deux approches est de sélectionner les caractéristiques les plus pertinentes, et ainsi d'accélérer l'apprentissage d'un modèle:
 - Pour la régularisation L_1 : [31] elle consiste à multiplier une constante de régularisation à la somme des poids en valeur absolue.
 - Pour la régularisation L_2 : [31] il s'agit de multiplier une constante de régularisation à la somme des poids au carré.

- Exploiter des techniques de dropout [32]: cette méthode, illustrée à la Figure 1.12, est couramment utilisée pour traiter le problème de sur-ajustement; elle consiste à désactiver de façon temporaire et aléatoire certains neurones dans le réseau et leurs connexions entrantes et sortantes.

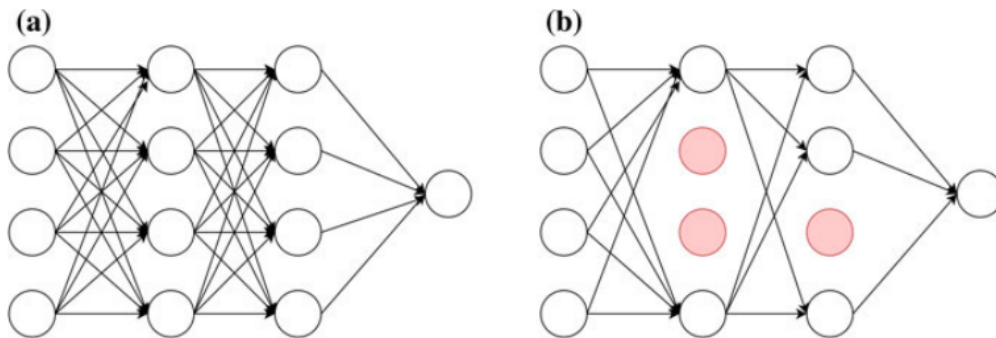


Figure 1.12: Illustration de la technique de Dropout : désactivation aléatoire de neurones

Early stopping

Pour éviter le sur-ajustement, le processus le plus élémentaire consiste à interrompre l'algorithme [33] quand la fonction de perte sur le jeu d'apprentissage commence à diminuer, et commence à croître sur le jeu de données de validation. Plusieurs approches ont été proposées pour permettre de déterminer facilement le seuil d'arrêt pendant la phase d'apprentissage comme la méthode de validation croisée (cross validation) [34] [35] .

1.5 Réseau de neurones convolutifs

Comme nous l'avons vu plus haut, il existe plusieurs types de réseaux de neurones: convolutifs, récurrents. Dans cette partie, nous allons plus particulièrement nous intéresser aux réseaux convolutifs qui seront utilisés tout au long de notre travail de thèse. Dans un premier temps, le principe général donnant l'architecture de ce type de réseau sera décrit en détail. Dans un second temps, nous montrerons comment l'appliquer sur des problèmes simples liés au domaine de l'image pour illustrer les possibilités d'une telle approche.

1.5.1 Architecture de base des Réseau de neurones convolutifs

Pour certains types de données, notamment pour les images, les perceptrons multi-couches classiques ne sont pas bien adaptés. En effet, ils sont définis pour des vecteurs en entrée. Pour les appliquer aux images, il faut transformer celles-ci en vecteurs en mettant bout à bout les lignes ou les colonnes des images. Ce processus aboutit au passage à la perte des informations spatiales contenues dans celles-ci, comme les formes des objets. De plus, les réseaux classiques sont constitués de couches pour lesquelles tous les neurones d'une couche (n) sont connectés à tous les neurones de la couche suivante ($n + 1$). Le nombre de poids affectés aux liens entre ces couches va augmenter de façon drastique lorsque le nombre de couches s'accroît. De ce fait, leur optimisation lors de la phase d'apprentissage va demander un temps de calcul extrêmement long. Les réseaux de neurones convolutifs (CNN) introduits par LeCun ont révolutionné le traitement des images et supprimé l'extraction des caractéristiques. Les CNN [12] [36] agissent directement sur les matrices, ou même sur les tenseurs pour les images à trois canaux de couleur RVB. Les CNN sont maintenant largement utilisés pour la classification et la segmentation d'images, la reconnaissance d'objets et de visages. Comme pour les réseaux classiques, les réseaux convolutifs sont organisés en couches placées en cascade. Toutefois, il est important de noter que toutes les couches ne jouent pas des rôles identiques. Ainsi, le choix de leur juxtaposition va être un facteur important dans la qualité du réseau obtenu. Les CNN font essentiellement apparaître trois types de couches, comme illustré à la Figure 1.13:

- Les couches convolutives,
- Les couches dites de « pooling »,
- Les couches entièrement connectées (fully connected),
- Les couches d'entrée et de sortie.

Nous allons maintenant voir plus en détail les caractéristiques spécifiques de chaque type de couche.

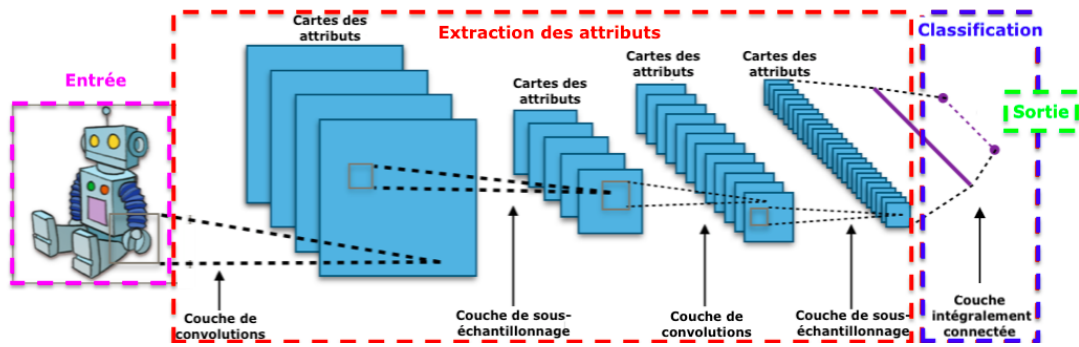


Figure 1.13: Illustration d'une architecture de réseau de type CNN.

1.5.2 Couche de convolution

La couche de convolution est le principal élément des réseaux de neurones convolutifs; cette couche a pour objectif d'extraire un ensemble de caractéristiques dans les images reçues en entrée. Pour y parvenir, on procède à une opération de convolution : La méthode est basée sur le concept de " déplacement " d'une fenêtre correspondant à une caractéristique sur l'image, et sur le calcul du produit de convolution entre la composante et chaque portion de l'image numérique.

Considérons deux fonctions notées f et g , le produit de convolution de ces deux fonctions est: $f * g$, ce produit est défini par l'équation suivante:

$$c(x) = f(x) * g(x) = \int_{-\infty}^{+\infty} f(t)g(x - t)dt \quad (1.24)$$

Dans un réseau de neurones convolutif, le premier élément f représente la donnée d'entrée (pixels d'une image), le deuxième paramètre équivalent au filtre de convolution, le résultat représente la donnée de sortie; ceci est appelée carte de caractéristiques, comme illustré à la Figure 1.14. L'entrée représente l'image sous forme d'un tenseur. Chaque tenseur est un tableau multidimensionnel contenant trois paramètres : hauteur, largeur et profondeur de l'image. Un filtre est représenté par une matrice de poids, ayant une hauteur h et une largeur k . Ces poids sont initialisés de manière aléatoire et sont ensuite mis à jour via la méthode de rétro-propagation du gradient.

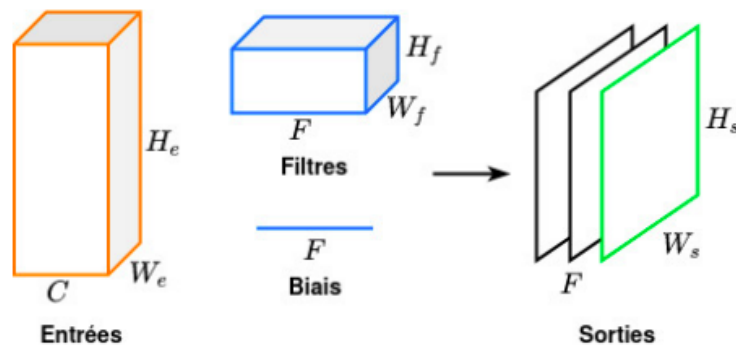


Figure 1.14: Convolution 2D et Génération de Carte Caractéristique

Suivant le type de problème à traiter, on réalise une convolution monodimensionnelle, bidimensionnelle ou tridimensionnelle. Nous allons détailler les opérations effectuées dans ces trois cas.

1.5.2.1 Convolution 1D

La convolution 1D s'applique lorsque le tenseur des données d'entrée se réduit à un vecteur. Il peut s'agir par exemple d'une mesure de courant ou de tension de fonction de la fréquence. La longueur du vecteur sera liée au nombre de points d'échantillonnage en fréquence dans cet exemple. Dans ce cas, on définit des filtres de convolutions qui seront des vecteurs de longueur généralement limitée à 3, 5 ou 7. Il est aisé de voir que l'opération de convolution revient à déplacer le filtre le long du vecteur avec un pas noté s pour « stride » fixe. La sortie obtenue sera alors un vecteur de longueur à priori plus importante que le vecteur d'entrée. Toutefois des opérations dites de zero padding permettent de conserver une longueur identique à celle du vecteur d'entrée. Considérons un vecteur d'entrée x de longueur L et un filtre de convolution f de longueur k (où k est généralement 3, 5 ou 7 comme mentionné). La sortie y de l'opération de convolution à une position i est donnée par :

$$y[i] = \sum_{j=0}^{k-1} x[i-j] \times f[j] \quad (1.25)$$

Lorsqu'on déplace le filtre le long du vecteur d'entrée avec un pas s (stride), la position i de la sortie est mise à jour en conséquence.

1.5.2.2 Convolution 2D

La convolution 2D s'applique lorsque le tenseur traduisant chaque donnée d'entrée est une matrice. Un exemple typique est celui des images représentées en niveau de gris. Dans ce cas, les filtres sont aussi des matrices typiquement de dimension 3*3, 5*5, 7*7. La convolution s'effectue cette fois dans les deux directions de l'espace de manière à balayer toute l'image avec un pas fixe noté s . Il permet ainsi d'avoir en sortie une image qui va traduire une caractéristique de l'image. Notons que les caractéristiques obtenues seront fonction du choix du filtre. Des exemples de structure de filtres seront illustrés par la suite.

Pour illustrer la convolution 2D en, considérons une image I de dimensions $M \times N$ et un filtre F de dimensions $k \times k$ (où k est typiquement 3, 5, 7, etc.). La valeur de l'image de sortie O à la position (i, j) est donnée par : La convolution 2D est donnée par :

$$O(i, j) = \sum_{m=0}^{k-1} \sum_{n=0}^{k-1} I(i + m, j + n) \times F(m, n) \quad (1.26)$$

1.5.2.3 Convolution 3D

L'objectif principal de cette opération, illustrée à la Figure 1.15, est la détection d'événements dans les vidéos et les images médicales 3D. Elle consiste à appliquer sur chaque entrée 3D un filtre qui se déplace dans 3 directions : de gauche à droite, de haut en bas et en profondeur. A la fin de cette opération, on produit une carte de caractéristiques de la forme 3D. Considérons un volume d'entrée V de dimensions $M \times N \times P$ et un filtre 3D F de dimensions $k \times l \times m$. La valeur du volume de sortie O à la position (i, j, k) est donnée par:

$$O(i, j, k) = \sum_{u=0}^{k-1} \sum_{v=0}^{l-1} \sum_{w=0}^{m-1} V(i - u, j - v, k - w) \times F(u, v, w) \quad (1.27)$$

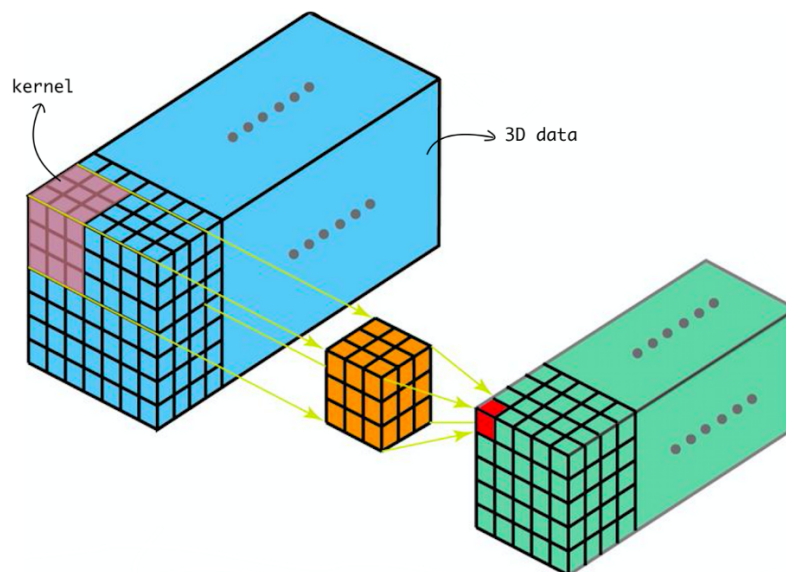


Figure 1.15: Convolution 3D

La couche de convolution se distingue par:

- La taille de noyau de convolution, qui représente les dimensions du filtre.
- Le nombre de filtres de convolution correspond au nombre de cartes de caractéristique, chaque sortie est représentée dans des tableaux multidimensionnels (tenseurs) où chaque tenseur est défini par une hauteur H , une largeur W , ainsi qu'un nombre de canaux sur chaque carte caractéristique qui correspondent au nombre de couleurs (RGB).
- Le stride s ou le pas de déplacement indique le nombre de pas de déplacement à chaque étape de la convolution., à titre d'exemple, lorsque le stride est égal à 1, nous déplaçons les filtres d'un pixel à la fois. Lorsque le pas est de 2, les filtres sont déplacés de 2 pixels lorsque nous les faisons glisser. Cela produit des volumes de sortie plus petits dans l'espace.
- Padding p , c'est un hyperparamètre qui permet de contrôler la taille d'entrée afin que la largeur et la hauteur de l'entrée et de la sortie soient identiques. Pour maintenir la dimension de la sortie comme celle de l'entrée, des pixels sont ajoutés autour de l'image d'entrée.

1.5.3 Type de convolution

On distingue divers types de convolutions:

1.5.3.1 Convolution classique

L'opération principale consiste à réaliser un produit scalaire entre deux matrices. La première est une matrice d'entrée qui représente une portion du champ récepteur, tandis que la seconde est un noyau représentant l'ensemble des paramètres apprenables (ou poids), comme illustré à la Figure 1.16. Ce noyau est généralement de taille plus petite que la matrice d'entrée. Quand l'image est composée de trois canaux (comme dans le cas des images RGB), le noyau possède une hauteur et une largeur spatialement réduites, mais il s'étend sur les trois canaux de profondeur. L'opération principale consiste à faire glisser ce noyau sur la hauteur et la largeur de l'image d'entrée, se déplaçant de gauche à droite puis de haut en bas, avec un pas de déplacement appelé "stride" (s). On obtient ainsi une matrice dimensionnelle de l'image appelée carte de caractéristiques (Feature maps or activation map) qui donne la réponse du noyau pour chaque position spatiale dans l'image.

Considérons, une image avec une dimension $X * X * H$ et H_s correspond au nombre de noyaux avec une taille spatiale de F avec un pas de glissement (stride) S et une quantité de remplissage P , la taille du volume de chaque carte caractéristique peut s'exprimer par l'équation suivante:

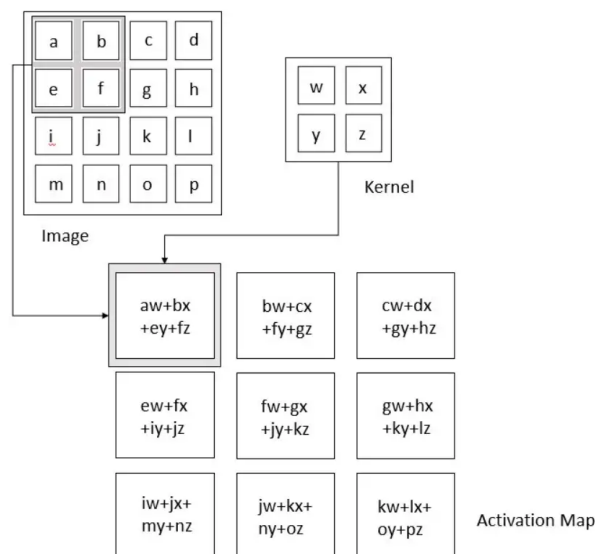


Figure 1.16: Opération de convolution [31]

$$C(i, j) = (I * F)(i, j) = \sum_{y=0}^{k-1} \sum_{x=0}^{k-1} I(s.i + x, s.j + y)F(x, y) \quad (1.28)$$

La convolution en vision par ordinateur est basée sur trois principes fondamentaux:

l'interaction clairsemée (ou sparse interaction), le partage de paramètres, et la représentation équivariante.

- **Interaction clairsemée (Sparse Interaction):** Dans un réseau de neurones entièrement connecté, comme le Multi-perceptron (MLP), chaque unité de sortie est liée à toutes les unités d'entrée. À l'inverse, dans un réseau convolutif, le nombre d'interactions est réduit. Ceci est dû au fait que le noyau de convolution est souvent bien plus petit que l'ensemble des données d'entrée. Cette caractéristique engendre une diminution du nombre de paramètres à calculer et à stocker, optimisant ainsi l'efficacité computationnelle et la performance statistique du modèle.
- **Partage des paramètres:** Lorsqu'une carte de caractéristiques est produite à partir de la convolution d'un noyau avec des données d'entrée, toutes les unités sur un même plan de la couche convolutive partagent les mêmes poids. Ainsi, on parle de partage de paramètres. Cette approche réduit considérablement le temps nécessaire à l'apprentissage, en grande partie grâce à la réduction du nombre de mises à jour des poids durant la rétropropagation.
- **Représentation équivariante:** Lorsqu'une entrée subit une transformation (comme une rotation ou une translation), sa représentation dans le réseau change de manière prévisible et cohérente. On parle alors de représentation équivariante. Cette propriété assure que les objets, quelle que soit leur position ou orientation, sont traités de manière uniforme par le réseau, garantissant une meilleure robustesse face à diverses transformations de l'entrée.

1.5.4 Couche de pooling

Les couches de mise en commun sont une des composantes des réseaux neuronaux convolutifs. Lorsqu'une couche de convolution est chargée d'extraire les caractéristiques des images, une couche de mise en commun consolide les caractéristiques apprises par les CNN. Leur but est de réduire progressivement la dimension spatiale de la carte caractéristiques afin de minimiser le nombre de paramètres et de calculs dans le réseau et de préserver les informations les plus pertinentes. Les couches de mise en commun permettent une invariance translationnelle qui permet au CNN d'être invariant aux translations, c'est-à-dire que même si l'entrée du CNN est transposée, le CNN sera toujours capable de reconnaître les caractéristiques de l'entrée. Une couche de mise en commun est ajoutée après les couches de convolution, elle permet de sous-échantillonner la carte caractéristique en appliquant un filtre de dimension $n \times n$

avec un certain pas prédéfini tout en calculant le maximum ou la moyenne de l'entrée. Il existe deux type de mise en commun [37] couramment utilisé dans le développement:

- **La mise en commun maximale maxpooling**: cette technique permet de sélectionner la valeur maximale de chaque groupe, cette méthode conserve les caractéristiques les plus pertinentes de la carte d'activation, avec une image plus claire, et moins bruité que l'image originale. La mise en commun maximale est couramment utilisée car elle préserve les caractéristiques détectées. Pour illustrer le concept de mise en commun maximale (ou « max pooling »), on peut considérer une fenêtre (ou filtre) de dimensions $n \times m$ que l'on déplace sur la carte d'activations. Si X est la matrice représentant la carte d'activations (où chaque élément de la matrice correspond à l'activation d'un neurone dans une position spécifique après la convolution et l'application de la fonction d'activation), et Y est la matrice résultant de l'opération de max pooling, alors chaque élément y_{ij} de Y est donné par :

$$y_{ij} = \max\{x_{kl} | k \in [i, i + n - 1], l \in [j, j + m - 1]\}$$

où x_{kl} est un élément de X et n et m sont les dimensions de la fenêtre de pooling. Cette équation implique que chaque élément de Y est le maximum des valeurs correspondantes dans une fenêtre $n \times m$ de X .

- **La mise en commun moyenne (Average-Pooling)**: cette opération permet de calculer la valeur moyenne de chaque segment de la carte des caractéristiques tout en conservant les caractéristiques fondamentales de l'image. Pour illustrer le concept de mise en commun moyenne (ou "Average Pooling"), considérons une fenêtre (ou filtre) de dimensions $n \times m$ que l'on déplace sur la carte des caractéristiques. À chaque emplacement de la fenêtre, on calcule la moyenne des $n \times m$ valeurs contenues dans cette fenêtre.

Si X est la matrice représentant la carte des caractéristiques et Y est la matrice résultant de l'opération de mise en commun moyenne, alors chaque élément y_{ij} de Y est donné par :

$$y_{ij} = \frac{1}{nm} \sum_{k \in [i, i + n - 1], l \in [j, j + m - 1]} x_{kl} \quad (1.29)$$

où x_{kl} est un élément de X et n et m sont les dimensions de la fenêtre de pooling. Cette équation signifie que chaque élément de Y est la moyenne des valeurs correspondantes dans une fenêtre $n \times m$ de X .

La Figure 1.17 illustre la mise en commun maximale (max pooling) et la mise en

commun moyenne (average pooling), où une carte d'activation de taille 7×7 est réduite avec des strides de 1 et 2.

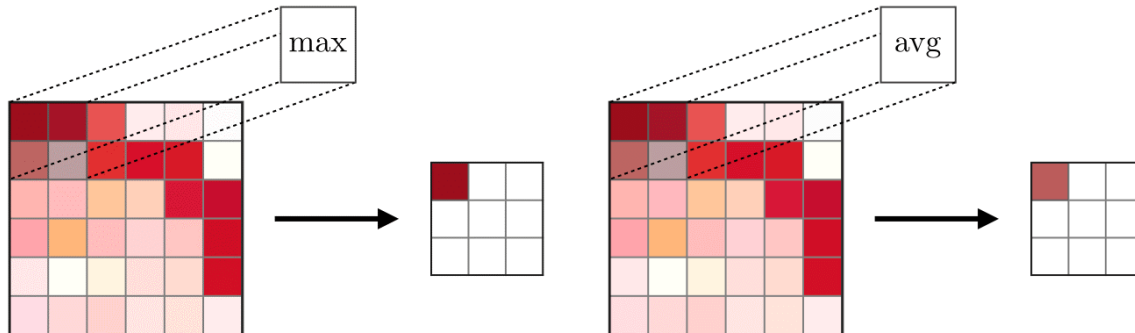


Figure 1.17: À gauche, exemple de max-pooling et à droite, exemple de average-pooling, d'une carte d'activation de taille 7×7 avec des strides de 1 et 2.

Afin de réduire la dimension spatiale des cartes d'activation, il est souhaitable d'utiliser un pooling avec un noyau de taille 2 x 2 et un pas de 2, un autre avantage important de la couche de mise en commun consiste à augmenter la taille du champ respectif tout en limitant la taille spatiale de la couche. La réduction de la dimension spatiale permet par ailleurs de diminuer les besoins en calculs, optimisant ainsi les performances du modèle. Une information spatiale limitée entraîne également une réduction du nombre de paramètres, ce qui évite les risques de surajustement.

Chaque couche de pooling nécessite un tableau multidimensionnel de dimension $W_1 \times H_1 \times P_{in}$, un noyau de taille N et un pas de déplacement qui représente le stride (s), pour chaque opération de mise en commun produit un tenseur de dimension $W_2 \times H_2 \times P_{out}$ ou

$$W_2 = \frac{W_1 - N}{s} + 1 \quad (1.30)$$

$$H_2 = \frac{H_1 - N}{s} + 1 \quad (1.31)$$

$$P_{in} = P_{out} \quad (1.32)$$

1.5.5 Couche complètement connectée

La couche entièrement connectée se trouve généralement à la fin d'un réseau de convolution et fonctionne de la même manière qu'un réseau de neurones

de type "feed-forward" multi-perceptron. Avant cette couche, il y a la couche d'aplatissement qui transforme les données en un tenseur unidimensionnel. Cette étape regroupe toutes les caractéristiques de la dernière couche spatiale et les relie à chaque unité de la première couche entièrement connectée. Selon la complexité de la tâche à accomplir, des couches denses supplémentaires peuvent être ajoutées pour améliorer la puissance de calcul et la performance du réseau. La couche de sortie d'un réseau neuronal convolutif est élaborée en fonction de son domaine d'application. La couche de sortie est totalement connectée à chaque neurone de l'avant-dernière couche, et est associée à un poids. En fonction de la nature de l'application : régression ou classification nous utilisons la fonction d'activation adéquate (softmax, linéaire, logistique).

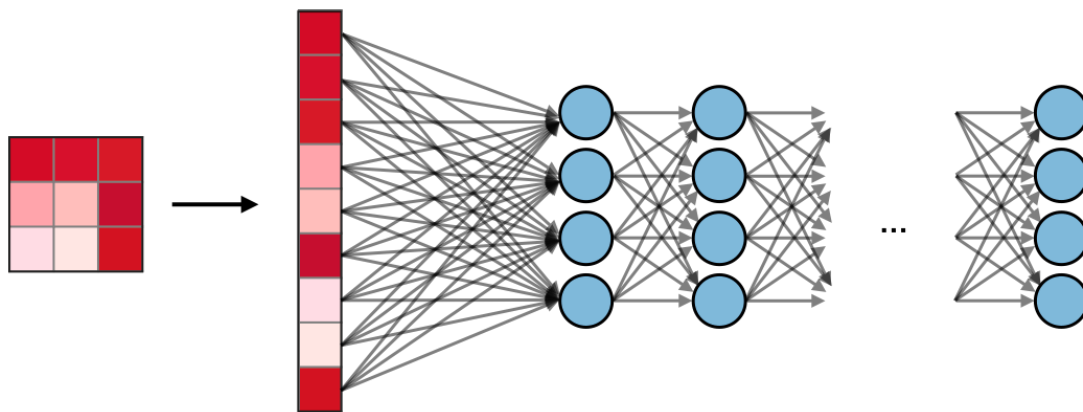


Figure 1.18: Couche complètement connecté

1.5.6 Paramètres spécifiques aux réseaux convolutifs

Formatage des données

Le formatage des données d'entrée dans un réseau de neurones convolutifs représente un ensemble d'étapes cruciales pour le bon fonctionnement du réseau convolutif. Cela inclut la sélection adéquate de la taille des données d'entrée, la normalisation des valeurs, le redimensionnement cohérent de toutes les images, le choix des couleurs du canal ainsi qu'un pré-traitement si cela s'avère nécessaire.

- Dimension des données d'entrée chaque entrée dans un réseau de neurones convolutifs est représenté sous forme de matrices 2D ou bien 3D. Généralement chaque image couleurs est représentée par trois canaux rouge, vert et bleu (RGB) où chaque canal est une matrice 2D. Le dimensionnement des données d'entrée 1D nécessite une séquence de valeur ayant la même

longueur car toutes les opérations de convolution nécessitent des filtres de taille constante

- *Normalisation des données* [29] [12]: Pour faciliter la convergence rapide d'un réseau de neurones à convolution, il est indispensable de fixer les valeurs de chaque pixel de l'image dans une plage commune de valeur, à titre d'exemple [0, 1] ou [-1, 1] de manière à ne pas privilégier une donnée par rapport à une autre.
- *Prétraitement des données*: Cette opération consiste à formater l'ensemble des données d'entrée pour améliorer la généralisation du modèle, il s'agit généralement d'utiliser des techniques comme : l'augmentation des données artificielles, détection des bords de l'image, la normalisation, centrage ou bien la standardisation des images.

Optimiseur

La descente de gradient est l'algorithme fondamental et le tout premier qui a été employé pour minimiser la fonction de perte d'un réseau de neurones en ajustant le taux d'apprentissage afin de converger vers des minima locaux. Cette méthode, bien que simple à implémenter, présente plusieurs limitations. Son fonctionnement repose sur le calcul du gradient en utilisant l'intégralité des données, suivi d'une mise à jour des poids du modèle. Cependant, cette première étape peut engendrer un ralentissement du processus d'entraînement. De plus, chaque itération requiert une quantité significative de mémoire, ce qui en fait une approche gourmande en ressources.

Descente de gradient stochastique

La descente de gradient représente un algorithme d'optimisation extrêmement répandu, néanmoins il existe diverses déclinaisons adoptées pour le perfectionnement des réseaux neuronaux. Parmi ces déclinaisons figure la version stochastique [38] où la descente de gradient stochastique est classiquement utilisée. Considérons J comme la fonction de coût à minimiser en fonction d'un vecteur de paramètres w .

$$J(w) = \frac{1}{m} \sum_{i=1}^m D_i(w) \quad (1.33)$$

En prenant en compte D_i comme étant la i -ème observation sélectionnée à

partir de l'ensemble d'apprentissage, au lieu de calculer le gradient de $J(w)$ en incorporant toutes les observations de D_i pour la modification des paramètres w , ce gradient est approximé en se basant sur une seule observation D_i .

À chaque itération, l'ensemble des poids w et des biais sont ajustés en suivant la formule qui suit :

$$w_{k+1} = w_k - \alpha \nabla D_i(w_k) \quad (1.34)$$

Cette méthode implique de calculer le gradient pour seulement une entrée à la fois dans la base d'entraînement, ce qui réduit grandement le coût par itération. Elle est particulièrement efficace lorsque la base de données est trop grande pour être stockée en mémoire ou n'est pas complètement accessible. De plus, la descente de gradient stochastique permet d'approcher plus rapidement la solution optimale grâce à des itérations plus courtes. Cependant, cette estimation du gradient pour une seule observation signifie qu'il faut effectuer plus d'itérations avec un taux d'apprentissage plus faible. Pour un bon équilibre, on peut considérer un groupe de données (appelé lot) plutôt qu'une seule observation, généralement de quelques dizaines à quelques centaines. Cela permet d'approximer le gradient de manière plus précise, ce qui facilite l'utilisation d'un taux d'apprentissage approprié pour converger rapidement.

Moments

Il existe d'autres approches pour accélérer la convergence de la descente de gradient, et l'une de ces méthodes est appelée "Momentum" [39] [40]. Cette méthode est inspirée de l'inertie d'un objet en mouvement. Le fonctionnement de la descente de gradient basée sur le momentum consiste à conserver une forme d'inertie tout en enregistrant les modifications antérieures pour calculer le déplacement de la fonction.

$$w_{k+1} = w_k - \alpha \Delta w \quad (1.35)$$

$$\Delta w_k = \gamma \Delta w_{k-1} + \frac{1}{b} \sum_{i \in X_t} \nabla D_i(w_k) \quad (1.36)$$

En utilisant le lot de données D_i à l'itération t , avec γ appartenant à l'intervalle ouvert $(0, 1]$, plus γ se rapproche de 1, plus l'exploration de l'espace des paramètres possède un moment significatif.

Les méthodes de second ordre

Les méthodes d'optimisation utilisant un gradient peuvent être divisées en deux approches distinctes. La première est fondée sur des approches de premier ordre, comme la descente de gradient classique. Ces méthodes se concentrent généralement sur l'ajustement des paramètres d'une fonction en utilisant sa dérivée première. D'un autre côté, il existe également des méthodes dites du second ordre, qui reposent sur une approximation du second ordre de la fonction à minimiser.

Il existe des algorithmes comme:

- L'algorithme BFGS (Broyden-Fletcher-Goldfarb-Shano) [41]
- L'algorithme DFP (Davidon-Fletcher-Powell) [42]

Autres techniques d'optimization

Pour accélérer la convergence du gradient, diverses approches ont été suggérées pour les réseaux neuronaux, telles que AdaGrad [43], Adadelta [44], RMSProp, ou Adam [45]. Parmi elles, la méthode d'optimisation Adam [45] est largement préconisée. Cet algorithme repose sur la méthode stochastique de descente de gradient qui utilise à la fois l'estimation adaptative des moments de premier ordre et de second ordre.

1.6 Architectures avancées

La partie suivante de cette thèse présentera les réseaux de neurones convolutifs utilisés:

1.6.1 VGG

Le réseau de neurones VGG16, connu sous le nom de Visual Geometry Group (VGG) et développé par l'Université d'Oxford [46] [47], a été conçu pour aborder les défis posés par les architectures de réseaux profonds. L'idée maîtresse derrière VGG16 est d'utiliser des filtres de petite taille (3x3) mais en augmentant la profondeur du réseau. Ceci permet d'avoir une meilleure représentation des caractéristiques tout en conservant une structure computationnelle efficace.

Contrairement aux architectures plus légères avec des filtres de différentes tailles, VGG16 maintient une uniformité en utilisant seulement des filtres de convolution

3x3 suivis de ReLU comme fonction d'activation. Le réseau suit généralement cette architecture : une série de couches de convolution, suivie d'une couche de maxpooling pour réduire les dimensions spatiales. Ceci est répété plusieurs fois pour capturer des caractéristiques à différentes échelles. La caractéristique distinctive de VGG16 est sa profondeur : il contient 16 couches pondérées, d'où son nom.

En ce qui concerne la structure détaillée, VGG16 commence avec deux couches de convolution suivies d'un pooling max, ce schéma est répété plusieurs fois avec une augmentation progressive du nombre de filtres de convolution. Finalement, après toutes les couches de convolution et de pooling, le réseau se termine par trois couches entièrement connectées. La dernière couche est une couche softmax qui classe les entrées dans l'une des 1000 classes (comme c'est couramment effectué pour la classification ImageNet [48]).

Le réseau VGG16 peut être décrit de la manière suivante:

$$\left\{ \begin{array}{l} x_0 = e \\ x_1 = f_{c1,w}(x_0) \\ x_2 = f_{p,w}(x_1) \\ \vdots \\ x_{2i+1} = f_{ci+1,w}(x_{2i}) \quad \text{avec } i = 1, \dots, 8 \\ x_{2i+2} = f_{p,w}(x_{2i+1}) \quad \text{avec } i = 1, \dots, 4 \\ \vdots \\ x_{13} = f_{fc1,w}(x_{12}) \\ x_{14} = f_{fc2,w}(x_{13}) \\ x_{15} = f_{fc3,w}(x_{14}) \end{array} \right. \quad (1.37)$$

Où:

- e : représente les données d'entrée du réseau.
- w : est le vecteur de poids sur lequel chaque couche dépend.
- $f_{c1,w}, f_{c2,w}, \dots$: sont les fonctions exprimant les couches de convolution de VGG16. Chaque fonction peut représenter plusieurs couches de convolution consécutives, toutes avec des filtres 3x3.
- $f_{p,w}$: représente la couche de pooling (généralement max pooling) qui réduit la taille spatiale de l'entrée de moitié.
- $f_{fc1,w}, f_{fc2,w}, f_{fc3,w}$: représentent les fonctions des couches complètement

connectées de VGG16, avec respectivement 4096, où 4096 et un nombre de dimensions de sortie équivalent au nombre de classes.

1.6.2 Resnet50

Le ResNet-50 est une variante du modèle ResNet (Residual Network), qui a été introduit par Kaiming [49]. ResNet a été conçu pour résoudre le problème de la disparition du gradient qui se pose lors de la formation de réseaux de neurones profonds. L'idée principale derrière cette architecture est d'introduire des "connexions résiduelles" (ou "sauts de connexion") qui permettent à l'information de contourner certaines couches.

Soit le bloc résiduel représenté par l'équation:

$$y_m = F_m(y_{m-1}) + y_{m-1} \quad (1.38)$$

Où:

- y_m est la sortie du m -ème bloc résiduel.
- F_m est une fonction englobant toutes les couches du bloc résiduel, comme les trois couches de convolution et les deux couches de ReLU illustrées dans la Figure x.

À partir de l'équation x, le réseau ResNet-50 peut être représenté comme :

$$\left\{ \begin{array}{l} x_0 = e \\ x_1 = f_{c,w}(x_0) \\ x_2 = f_{p,w}(x_1) \\ \vdots \\ x_{t+1} = F_{1,w}(x_t) + x_t, \quad t = 2, \dots, 7 \\ \vdots \\ x_{t+1} = F_{2,w}(x_t) + x_t, \quad t = 8, \dots, 20 \\ \vdots \\ x_{t+1} = F_{3,w}(x_t) + x_t, \quad t = 21, \dots, 38 \\ \vdots \\ x_{t+1} = F_{4,w}(x_t) + x_t, \quad t = 39, \dots, 49 \\ x_{50} = f_{s,w}(x_{49}) \end{array} \right. \quad (1.39)$$

Où :

- e : données d'entrée du réseau.
- w : vecteur de poids associé à chaque couche.
- $f_{c,w}$: fonction représentant la couche de convolution conv1.
- $f_{p,w}$: fonction représentant la couche de max pooling.
- $f_{s,w}$: fonction représentant la couche de sortie du réseau.
- $F_{l,w}$: fonction représentant le bloc résiduel de la l -ième couche de convolution.

1.7 Conclusion

Dans ce chapitre, nous avons posé les bases de l'apprentissage automatique en mettant l'accent sur différents concepts essentiels. Nous avons exploré les diverses métriques permettant d'évaluer nos modèles développés. Nous nous sommes ensuite concentrés sur les réseaux de neurones profonds, abordant leur principe de fonctionnement et leur processus d'apprentissage. Nous avons détaillé les étapes nécessaires pour comprendre les réseaux de neurones à convolution et avons examiné leurs différentes couches ainsi que les paramètres spécifiques à ces réseaux convolutifs. Nous avons également présenté deux architectures avancées : VGG16 et ResNet-50, qui seront utilisées dans le cadre de cette thèse. Forts de ces fondamentaux, nous aborderons dans les chapitres suivants des applications concrètes liées à notre domaine d'étude, notamment le Radar GPR et le diagnostic des défauts sur câbles.

2

Etat de l'art sur la détection et la localisation de défauts dans les câbles par réflectométrie

Sommaire

2.1	Introduction	65
2.2	Domaines d'application des câbles	66
2.3	Types de câbles électriques	68
2.4	Classification des Anomalies Présentes dans les câbles électriques	72
2.5	Etat de l'art sur le diagnostic des défauts dans les câbles	75
2.6	Concept de réflectométrie et théorie des lignes	76
2.7	Modélisation d'une ligne	78
2.8	Méthodes basées sur la réflectométrie	84
2.9	Réflectométrie classique	84
2.10	Conclusion	89

2.1 Introduction

Avec l’accroissement du nombre de systèmes électroniques, le nombre de calculateurs se trouve lui aussi être en forte augmentation. Cette problématique se rencontre notamment dans le domaine des transports comme l’aéronautique et le domaine automobile. Ainsi, il est essentiel de concevoir des liaisons possédant une forte fiabilité pour assurer la communication d’une commande jusqu’au calculateur et du calculateur vers l’actionneur pour des raisons de sécurité. On rencontre aussi des liaisons par câble devant posséder une forte fiabilité dans les bâtiments de production d’énergie pour ne citer que ceux-ci où la sécurité est un enjeu crucial. Si, comme on vient de le voir, le nombre de câbles augmente sans cesse, un autre paramètre est lui aussi en croissance constante, il s’agit de la longueur des câbles. La Figure 2.1 montre la longueur cumulée des câbles dans différents domaines d’application. Compte tenu de cette remarque, il apparaît que la surveillance de la "bonne santé" du réseau est importante et devient de plus en plus difficile et longue à réaliser par des moyens classiques. Il est donc important de développer des méthodes appropriées permettant de réaliser cette opération.

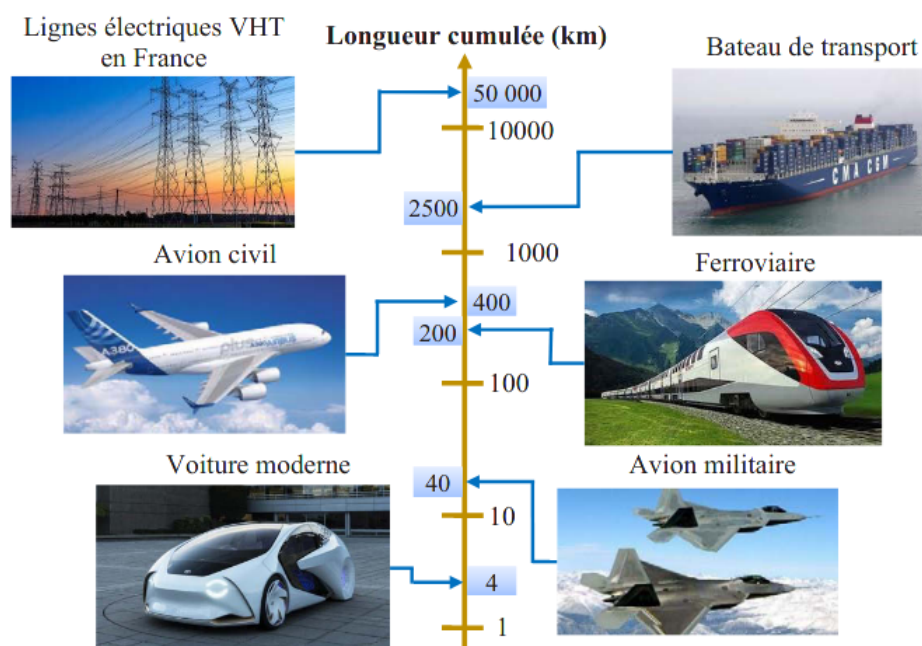


Figure 2.1: Longueurs cumulées de câbles dans les différentes applications (en km)

Pour répondre aux nouvelles exigences en matière de sécurité et de qualité de fonctionnement, il est impératif de mettre en œuvre des stratégies permettant

d’améliorer la qualité de service. Cela implique l’élaboration de techniques sophistiquées pour la surveillance de ces lignes de transmission d’électricité ou de signaux. En effet, dans divers secteurs tels que l’automobile et l’aéronautique, comme vu plus haut, la longueur cumulée des câbles électriques a connu une croissance exponentielle. A titre d’exemple, dans un avion de transport civil moderne tel l’A380, on trouve désormais jusqu’à 500 km de câbles électriques. Ainsi, l’adoption de méthodes de surveillance améliorées est d’une importance capitale pour maintenir l’efficacité de ces systèmes.

Le présent chapitre sera organisé comme suit : nous présenterons une technique de localisation des défauts, dans laquelle nous nous concentrerons uniquement sur les approches reposant sur la réflectométrie. Nous aborderons dans ce contexte la notion de ligne de transmission et une technique de modélisation, afin d’acquérir une bonne compréhension du principe de ce type d’approche. Par la suite, nous nous pencherons sur les méthodes et les types de signaux injectés, à la fois dans le domaine temporel et dans le domaine fréquentiel. Nous expliquerons son mode de fonctionnement, ainsi que ses avantages et ses limites. Dans cette thèse, nous nous focaliserons exclusivement sur l’utilisation de la réflectométrie dans le domaine temporel.

2.2 Domaines d’application des câbles

Avionique et espace

L’industrie aéronautique fournit une illustration significative de la complexité des réseaux de câblage, soulignant l’importance cruciale de la détection et du diagnostic précis des défauts dans les faisceaux de câbles. Les résultats d’une enquête de l’USAF (Force Aérienne des États-Unis) ont révélé que les défaillances dans les systèmes de câblage peuvent mener à des incidents sérieux et même à des catastrophes tragiques.

Pour identifier et classer les différents types de dysfonctionnements pouvant survenir dans les réseaux de câbles, de nombreuses études ont été entreprises par la FAA. Ces recherches ont mis en lumière diverses anomalies telles que les connecteurs endommagés ou cassés, la corrosion des câbles, des câbles arrachés ou coupés, ainsi que les circuits courts ou ouverts [50] .

Des travaux de recherche de la NASA ont également révélé que les aéronefs de plus de 20 ans sont souvent confrontés à des problèmes de câblage. Ces

défaillances peuvent avoir des conséquences catastrophiques, menant à des accidents d’avions.

L’une des tragédies les plus marquantes imputable à un dysfonctionnement du réseau de câbles est l’accident du vol TWA 800 [51] [52] [53], qui a explosé en plein vol une dizaine de minutes après son décollage, emportant avec lui 230 passagers. Une autre catastrophe aérienne majeure, celle du vol 111 [52] de Swissair en 1998, a coûté la vie à 229 personnes et est considérée comme l’un des accidents les plus graves de l’histoire de l’aviation civile. D’après les enquêteurs, l’accident a été causé par un arc électrique provenant d’un revêtement de fil corrodé, qui a déclenché l’incendie d’une couche d’isolant inflammable dans le plafond de l’avion.

Ces exemples tragiques soulignent l’importance cruciale d’un diagnostic précis et d’une gestion efficace des problèmes de câblage dans l’industrie aéronautique. Il est indispensable d’adopter des mesures proactives pour éviter que de tels accidents ne se reproduisent.

Automobiles

L’évolution de l’industrie automobile s’est traduite par une augmentation significative de la complexité de ses systèmes électroniques. Cette évolution est en grande partie due à la technologie X-by-wire [54], dont l’objectif est de substituer les systèmes mécaniques et hydrauliques traditionnels par des alternatives électroniques plus modernes.

Dans le contexte de l’automatisation et de l’électrification croissante, une grande diversité de réseaux locaux est devenue nécessaire. Ces réseaux sont indispensables pour faciliter la communication entre les multiples unités de contrôle électronique (UCE) présentes dans un véhicule moderne.

Chaque composant interagit avec les autres par le biais de faisceaux de câbles, qui assurent la transmission des signaux électriques et permettent la circulation des données collectées. Face à cette complexité croissante, la qualité et la sécurité des faisceaux de câbles sont devenues une priorité. Il est désormais impératif d’assurer une surveillance rigoureuse et une maintenance proactive de ces éléments clés pour garantir le bon fonctionnement des véhicules modernes.

2.3 Types de câbles électriques

Un câble est essentiellement un regroupement de conducteurs parallèles, qui peuvent être blindés ou non, tous contenus dans un même revêtement. Au fil des années, divers types de câbles ont été conçus, chacun optimisé pour répondre à des exigences spécifiques associées à des domaines d’application particuliers comme l’aéronautique, l’automobile, les télécommunications, etc.

La sélection d’un câble dépend de plusieurs critères. Parmi ceux-ci figurent la nature du signal à transmettre (analogique ou numérique, énergie ou signal, basse ou haute fréquence, etc.), le domaine d’application visé, le débit de transmission recherché, et la longueur du réseau à couvrir.

Nous allons maintenant aborder les détails de trois grandes catégories de câbles: le câble coaxial [55], la paire torsadée [56] et la ligne de transmission bifilaire.

Le câble coaxial, souvent utilisé en télécommunications, est connu pour sa capacité à transporter des signaux de haute fréquence sur de longues distances avec un minimum de pertes de signal. Il se compose d’un conducteur central entouré d’un blindage, ce qui le rend résistant aux interférences électromagnétiques.

La paire torsadée, comme son nom l’indique, est composée de deux conducteurs enroulés l’un autour de l’autre. Ce design permet de réduire l’interférence électromagnétique entre les conducteurs de paires voisines dans une même gaine. Cette caractéristique fait de ce câble un choix courant pour la transmission de données dans les réseaux informatiques.

Enfin, la ligne de transmission bifilaire se compose de deux conducteurs parallèles; elle est couramment utilisée dans des applications de faible puissance et de haute fréquence, comme dans le domaine des antennes.

Ces trois types de câbles illustrent la diversité des solutions disponibles pour répondre aux exigences variées de transmission de signaux et d’énergie dans diverses industries et applications.

2.3.1 Le câble coaxial

Le câble coaxial est une ligne de transmission asymétrique, couramment utilisée pour la transmission de signaux à basse ou haute fréquence. Ce câble se compose de deux conducteurs, un central et un extérieur. Le conducteur central peut être constitué de plusieurs fils ou d’un seul (le matériau couramment utilisé est le cuivre, mais il peut aussi être étamé, argenté ou même constitué d’acier cuivré),

et est enveloppé par un matériau diélectrique, ou isolant. Cet isolant joue un rôle crucial car il protège non seulement les données contre le bruit et les interférences externes, mais permet aussi une transmission rapide sur de longues distances. Ce matériau diélectrique peut être enveloppé par une tresse conductrice (simple ou double) — une bande ou une feuille de cuivre ou d’aluminium peut aussi être enroulée autour — ou un tube fait de différents matériaux comme le cuivre nu, le cuivre annelé, le cuivre étamé, ou l’aluminium étamé. Enfin, le tout est protégé par une gaine extérieure isolante et protectrice. Plusieurs caractéristiques électriques définissent le câble coaxial : la fréquence de fonctionnement ou la fréquence de coupure du câble, les pertes en signal ou l’atténuation mesurée en décibels par mètre (dB/m), l’impédance caractéristique (généralement de 50 ou 75 ohms), la vitesse de propagation du signal et le retard linéaire, qui sont liés au matériau diélectrique utilisé. En somme, le câble coaxial est un composant robuste et fiable pour la transmission de signaux, notamment grâce à sa protection contre les interférences et sa capacité à maintenir la qualité du signal sur de longues distances.

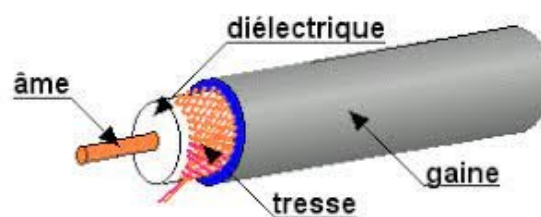


Figure 2.2: Description matérielle d'un câble coaxial (source: Wikipédia).

2.3.2 Paire torsadé

Une paire torsadée est une configuration de câblage qui se compose de deux conducteurs individuels enroulés ensemble en une configuration hélicoïdale. Cette disposition a un objectif primordial qui est de limiter la sensibilité aux interférences électromagnétiques ainsi que la diaphonie, c'est-à-dire le brouillage entre signaux de différentes paires de câbles lorsqu'elles sont regroupés dans un même faisceau de câbles. La conception de la paire torsadée aide à annuler les interférences électromagnétiques qui peuvent provenir de sources externes ou même des autres fils du faisceau. Le fait de torsader les paires permet d'annuler le flux de champ magnétique total induit dans les boucles constituées par chaque

torsade. En effet, d’une torsade à la suivante, ces flux donnent des courant en opposition de phase, ce qui contribue à une baisse du niveau du courant parasite induit en extrémité de celle-ci sur une paire.

Les câbles à paires torsadées peuvent se composer d’une seule paire, comme cela se voit couramment dans les lignes téléphoniques, ou de plusieurs paires. Les câbles comportant deux ou quatre paires sont couramment utilisés dans les réseaux informatiques, par exemple les câbles Ethernet. Notons que les câbles Ethernet modernes peuvent contenir jusqu’à huit paires torsadées dans une seule gaine.

En fonction des exigences du système, les paires torsadées peuvent être blindées ou non blindées. Le blindage, qui peut recouvrir chaque paire individuelle ou l’ensemble du câble, fournit une protection supplémentaire contre les interférences électromagnétiques.

En somme, les paires torsadées sont un choix à privilégier pour les transmissions de données et de voix, car elles offrent un équilibre entre le coût, la facilité d’installation et la protection contre les interférences.

- **La paire torsadée non blindée**, souvent désignée par son acronyme UTP pour Unshielded Twisted Pair, est un type de câble qui n’est pas enveloppé d’un blindage de protection contre les interférences électromagnétiques. Il s’agit de l’un des types de câblage les plus couramment utilisés dans les réseaux de télécommunication et de données en raison de son coût relativement faible et de sa facilité d’installation.
- **La paire torsadée blindée**, souvent abrégée STP pour Shielded Twisted Pair, est un type de câble de transmission de données qui comprend un blindage de protection contre les interférences électromagnétiques. Ce blindage est généralement constitué d’une feuille d’aluminium qui entoure toutes les paires de fils à l’intérieur du câble et la gaine extérieure isolante. Les interférences électromagnétiques peuvent provenir de diverses sources telles que des appareils électroniques avoisinants, des lignes d’énergie électrique, des moteurs, etc. Lorsqu’elles sont présentes dans un environnement où le câble STP est utilisé, elles peuvent créer des distorsions dans le signal transmis, pouvant conduire à des erreurs de données.
- **La paire torsadée super blindée** est une technologie de câblage avancée où chaque paire de fils est protégée individuellement par un écran en aluminium. Cette méthode de blindage offre une protection robuste contre les interférences électromagnétiques potentielles, renforçant ainsi l’intégrité des

données transmises.

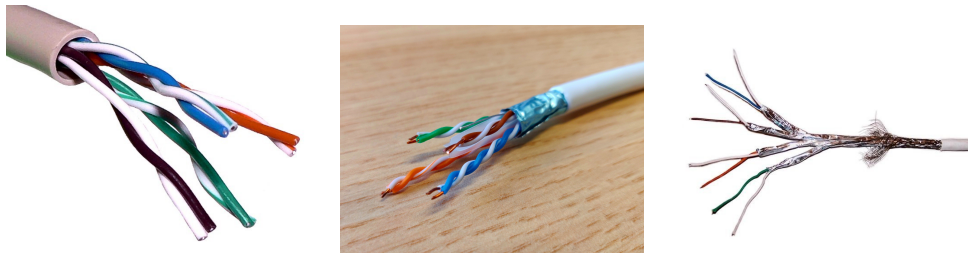


Figure 2.3: Figure montrant les trois types de paires torsadées: non blindées, blindées et super blindées (source : Wikipédia).

2.3.3 Câbles d’énergie

Les câbles d’énergie jouent un rôle fondamental dans la transmission de l’électricité à haute tension. Ils sont conçus pour véhiculer efficacement l’énergie, qu’elle soit sous forme de courant alternatif (AC) ou de courant continu (DC).

- Concernant le courant alternatif, il est généralement utilisé pour la distribution d’électricité sur de longues distances, notamment dans les réseaux électriques résidentiels et commerciaux.
- D’autre part, le courant continu est couramment utilisé dans les appareils électroniques ainsi que dans certaines applications nécessitant une alimentation stable en courant électrique. Cependant, pour les transmissions longue distance, en particulier dans les réseaux électriques, le courant alternatif est généralement préféré en raison de sa capacité à réduire les pertes d’énergie sur de grandes distances.

Ces câbles d’énergie à haute tension assurent le transport de l’électricité depuis les centrales de production d’énergie vers les sous-stations de distribution, puis vers les consommateurs finaux.

Grâce à leur conception robuste, ces câbles peuvent supporter les hautes tensions nécessaires pour la transmission de l’électricité à grande échelle. De plus, ils sont construits pour résister à diverses conditions environnementales et pour minimiser les pertes d’énergie lors de la transmission, optimisant ainsi l’efficacité globale du système électrique.

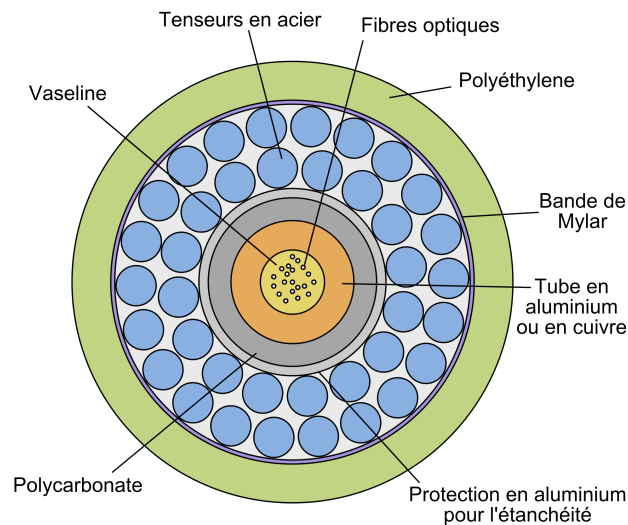


Figure 2.4: Coupe transversale d'un câble d'énergie (source wikipedia)

Après avoir présenté différents types de câbles ainsi que leurs utilisations spécifiques, nous aborderons à présent les principaux défauts auxquels un câble peut être sujet ainsi que les principales causes de leur occurrence.

2.4 Classification des Anomalies Présentes dans les câbles électriques

2.4.1 Défauts non francs

Au fil des années, les câbles subissent une dégradation naturelle qui les rend de plus en plus vulnérables, tout comme tout autre composant électrique ou électronique. Parmi les types de défauts que peuvent rencontrer les câbles, on retrouve les "défauts mous ou défauts non francs". D'après les constatations de la NASA, ces défauts représentent environ 45% des problèmes rencontrés dans les câbles utilisés en aéronautique. Les défauts non francs peuvent avoir différentes origines telles que : que le frottement excessif, la corrosion et les problèmes d'isolation. Bien que ces défauts ne conduisent pas directement à des incidents catastrophiques, ils peuvent engendrer des effets indésirables tels que des points chauds (zones de surchauffe) ou des défauts francs tels que les courts-circuits et les circuits ouverts.

La détection et la localisation précise de ces petits défauts jouent un rôle crucial

pour la maintenance prédictive. En anticipant ces problèmes potentiels, il est possible de planifier les interventions de maintenance de manière proactive, ce qui permet d’éviter les dysfonctionnements du système. Cette prévention s’avère essentielle, par exemple dans le domaine d’aéronautique notamment lorsque l’on considère que chaque heure d’immobilisation d’un avion au sol peut coûter environ 150 000 dollars.

Cependant, la détection des défauts non francs pose un réel défi technique. Ces défauts se traduisent par des signaux de faible amplitude qui peuvent être noyés au milieu d’interférences ou des bruits électriques [57] [50], ce qui les rend difficiles à repérer. Il est donc primordial de développer des méthodes de détection avancées et suffisamment sensibles pour identifier ces problèmes à temps et assurer ainsi la fiabilité et la sécurité des systèmes.

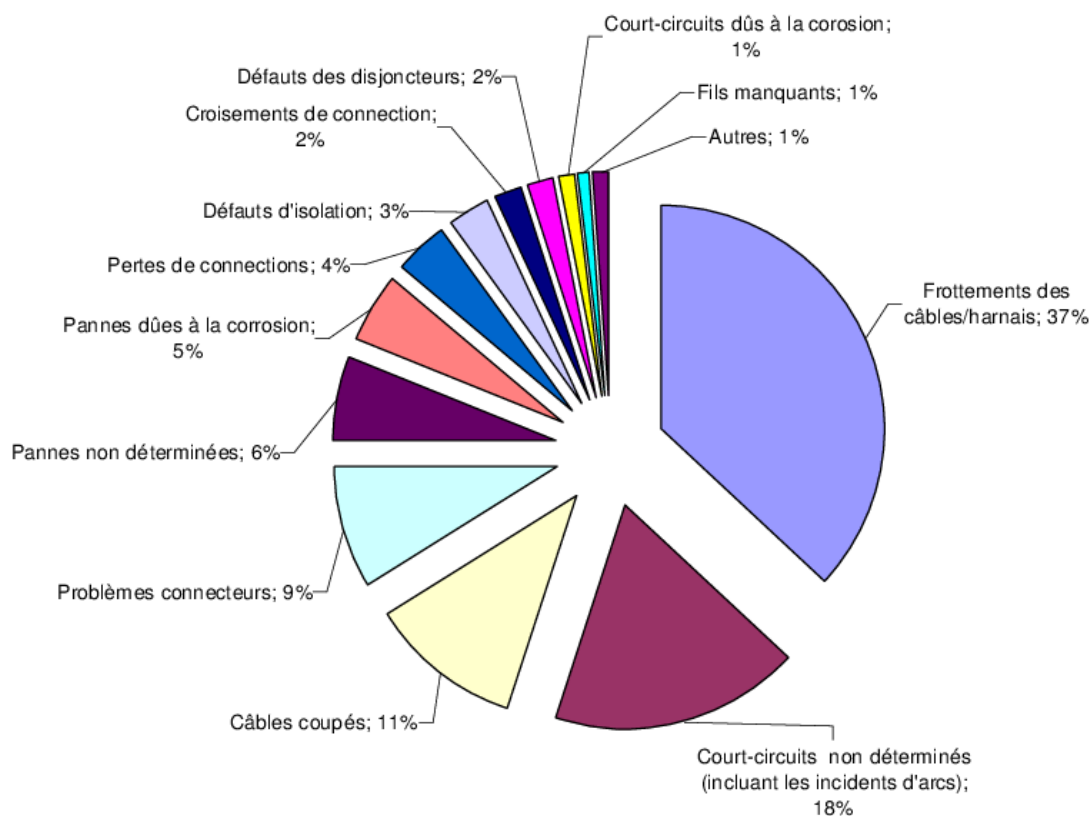


Figure 2.5: Types des défauts et leur répartition selon la base de données Navy Safety Center Hazardous Incident entre 1980 et 1999 [57].

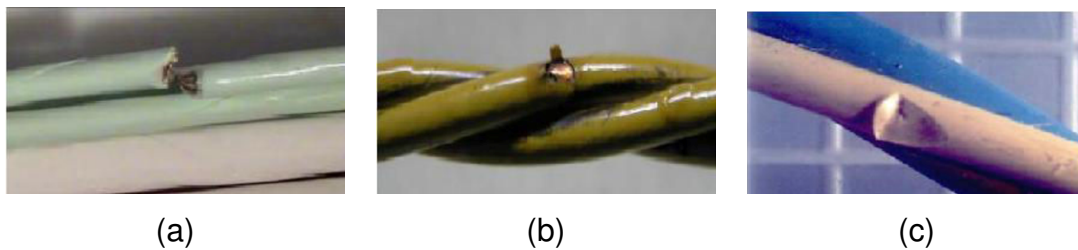


Figure 2.6: Exemples de défauts: (a) circuit-ouvert (défaut franc), (b) perte de l’isolant (défaut non franc) et (c) défaut d’écrasement (défaut non franc) [58].

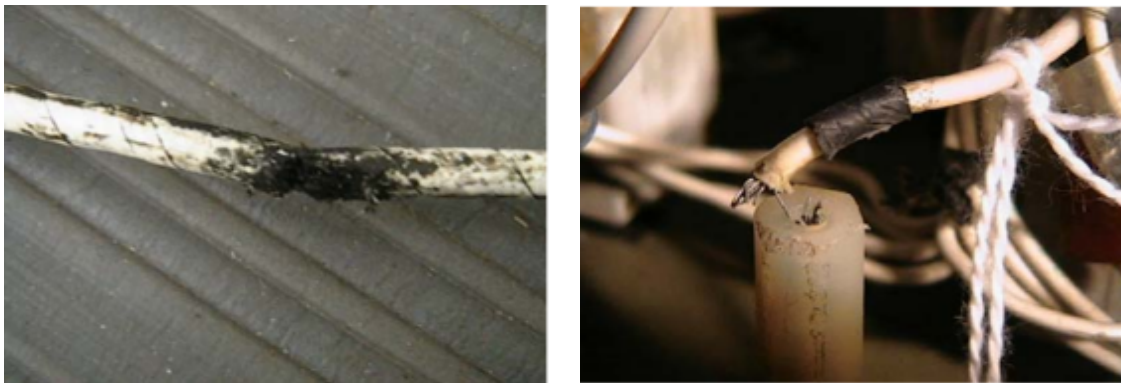
2.4.2 Défauts sévères

Les défauts francs sont des anomalies particulièrement préoccupantes, se caractérisant par une variation significative locale de l’impédance caractéristique du câbles qui peut être assimilée à un circuit ouvert ou à un court-circuit. Ils engendrent une interruption totale du flux d’énergie ou d’information dans le câble endommagé.

Le circuit ouvert survient lorsque l’un ou plusieurs conducteurs du câbles sont rompus suite à des dommages mécaniques ou des mouvements violents. Cette discontinuité empêche le passage du courant électrique, entraînant une perte de connectivité et de transmission des signaux. En revanche, le court-circuit survient lorsque deux ou plusieurs conducteurs se trouvent accidentellement en contact, présentant une faible résistance en raison d’une isolation endommagée. Cette situation crée un chemin de moindre résistance pour le courant, générant des courants intenses et potentiellement dangereux.

Les circuits ouverts et les court-circuits peuvent entraîner des accidents tragiques en cas de dysfonctionnement soudain du système auquel le câble est connecté. Ces situations peuvent causer des pannes critiques, la perte de fonction souhaitée au niveau du calculateur, des risques d’incendie ou encore des accidents mettant en danger la sécurité des personnes et des équipements.

Pour assurer la sécurité des systèmes, la détection précoce et la correction rapide de ces défauts francs sont primordiales. Les procédures de maintenance et de surveillance doivent être rigoureusement mises en place pour minimiser les risques liés à ces dégradations graves. En investissant dans des technologies de diagnostic avancées et en adoptant une approche proactive.



(a): Circuit ouvert

(b): Coupure d'un conducteur

Figure 2.7: Exemple de défauts sévères dans les câbles aéronautiques [59].

2.5 Etat de l’art sur le diagnostic des défauts dans les câbles

La réflectométrie est une technique avancée utilisée pour détecter et localiser les défauts dans les câbles électriques. Cette méthode est devenue d’une importance cruciale, notamment après l’accident tragique impliquant la navette spatiale Columbia STS-93, ou un court circuit dans les ordinateurs principaux a entraîné une explosion.

Suite à cet événement marquant, la NASA a entrepris des recherches approfondies pour améliorer les méthodes de test et de diagnostic des câbles afin d’éviter de futurs incidents. l’une des principales conclusions du rapport “Investigation of the challenger accident” [60] fut la nécessité de développer des méthodes plus efficaces pour détecter et localiser les pannes dans les câbles.

Une des premières méthodes de diagnostic mentionnées dans le rapport est l’inspection visuelle. Cette approche consiste à examiner visuellement les réseaux de fils afin de repérer toute anomalie évidente. Elle s’avère utilisée pour détecter les échauffements locaux des câbles et la dégradation de l’isolation. Cependant, elle peut se révéler limitée pour réparer les défauts internes aux câbles.

Pour surmonter cette limitation, les chercheurs ont exploré l’utilisation des rayons X pour la localisation des défauts. Cette technique permet d’inspecter les câbles de manière non destructive et d’identifier les problèmes au niveau des conducteurs internes ainsi qu’au niveau de l’isolation externe. Grâce à cette méthode, il est possible de repérer des défauts cachés sans altérer physiquement les câbles.

L’avènement de la réflectométrie a marqué un tournant majeur dans la détection des défauts dans les câbles. Cette méthode novatrice offre une approche globale pour évaluer la qualité d’un câble en mesurant les ondes réfléchies qui se propagent le long de celui-ci.

La réflectométrie est une technique avancée basée sur l’exploitation de la propagation de signaux haute fréquence. Son fonctionnement repose sur l’injection d’un signal dans un câble, suivi de la détection des ondes réfléchies aux différentes discontinuité d’impédance caractéristique présentes dans le câbles. Cette méthode permet d’analyser la façon dont les signaux se propagent le long du câble et d’identifier les variations d’impédance qui peuvent être dues à des défauts tels que des courts-circuits, fissure dans les câbles, ou d’autres anomalies. L’avantage de la réflectométrie réside dans sa capacité à localiser précisément ces défauts sur toute la longueur du câbles, offrant ainsi un outil puissant pour le diagnostic et la maintenance sur l’ensemble des câbles.

Pour appréhender pleinement le principe de la réflectométrie, il est essentiel de présenter préalablement la théorie des lignes qui en constitue le fondement théorique. Elle décrit le comportement des signaux électriques le long des câbles et permet de comprendre comment les ondes se propagent et interagissent avec les différentes caractéristiques des câbles, telles que l’impédance, et apporte une compréhension des phénomènes de propagation et du couplage du champ environnant avec un câble, ainsi que du couplage direct entre les conducteurs d’un toron. Une fois cette base établie, nous aborderons en détails les différentes techniques de réflectométrie temporelle.

2.6 Concept de réflectométrie et théorie des lignes

La réflectométrie [61] est une approche non destructive qui consiste à injecter le long du fil un signal de test spécifique caractérisé par une basse tension et une haute fréquence. Ce signal est inséré depuis l’extrémité du câble électrique, se propageant ainsi le long de sa trajectoire. Une partie de son énergie est ensuite réfléchi lorsque l’onde rencontre une jonction ou un défaut. Grâce à cette technique, il est possible d’analyser les réflexions engendrées par ces discontinuités [62] et d’établir un réflectogramme qui fournit des informations détaillées sur les caractéristiques du câble ainsi que sur les éventuels défauts. Ces informations incluent la localisation précise du défaut, son type et sa taille.

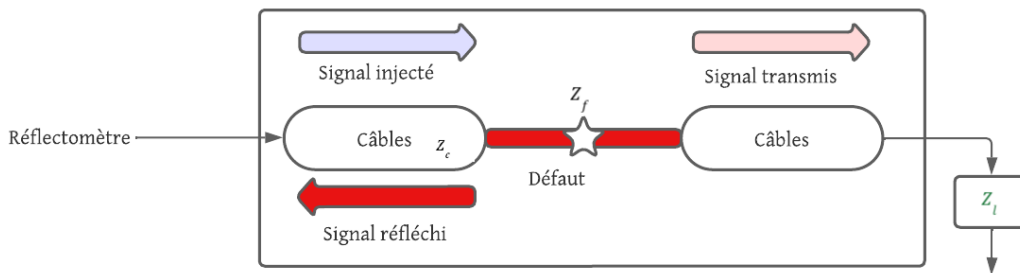


Figure 2.8: Principe de la réflectométrie

Le principe de base de la réflectométrie temporel est l'émission à une extrémité d'un réseau de câbles d'un signal $x_i(t)$, le réseau pouvant être constitué d'un seul câble ou de câbles interconnectés qui introduisent des réflexions sur la source liées aux discontinuités d'impédances au niveau des embranchements et des réflexions de forte amplitude intervenant aux extrémités libres (ou du moins désadaptées) des câbles. Cette mesure est tout d'abord réalisée sur un câble sain. Ainsi, si on définit un signal $x_c(t)$ traduisant la réponse impulsionnelle du canal de propagation, le signal $x_r(t)$ est alors la convolution de cette réponse et du signal incident, son expression est donnée par:

$$x_r(t) = x_i(t) * x_c(t) \quad (2.1)$$

Lorsqu'un défaut apparaît, des réflexions supplémentaires sont alors observées, dans ce cas, le signal réfléchi est modifié du fait du changement de la réponse impulsionnelle. Ainsi en effectuant des mesures régulières et en observant la différence du signal réfléchi à l'instant courant de celle obtenue à la récurrence précédente, on observe alors un signal non nul portant les caractéristiques du défaut. Ces caractéristiques sont : l'instant auquel revient l'impulsion qui est liée à sa position, l'amplitude du défaut liée au coefficient de réflexion et donc à l'impédance du défaut.

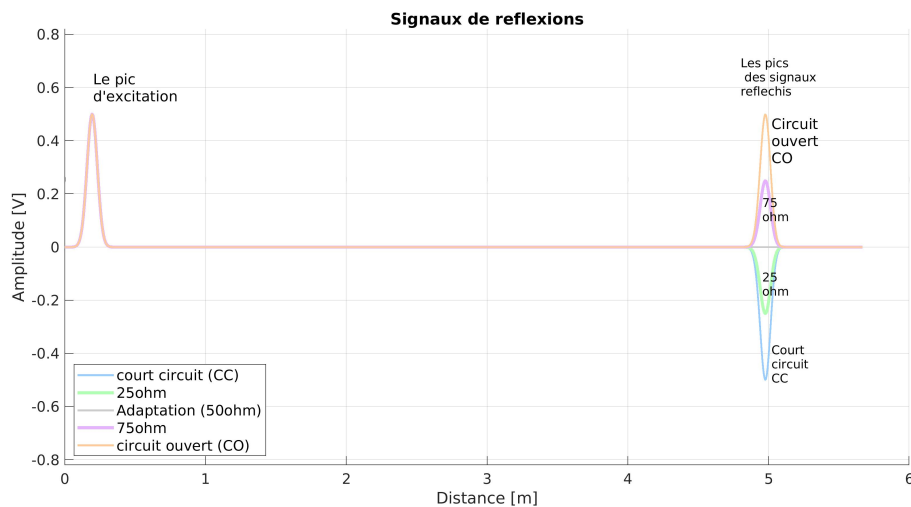


Figure 2.9: Réflectogrammes pour différentes valeurs de Z_l sur un câble coaxial d'impédance caractéristique $Z_c = 50 \Omega$

2.7 Modélisation d'une ligne

Dans cette section, nous rappelons brièvement les concepts liés aux lignes des transmission [63] l'objectif étant de préciser le processus de propagation d'une onde électromagnétique au sein d'une ligne de transmission et ensuite de déduire le mécanisme sous-jacent à la réflectométrie et d'être en mesure d'interpréter le tracé d'un réflectogramme.

2.7.1 Régime temporel quelconque

Afin de modéliser une ligne de transmission constituée de deux conducteurs (un conducteur aller et un retour), nous discrétisons celle-ci en tronçons élémentaires de longueur très inférieure à la longueur d'onde du signal se propageant sur celle-ci ($\lambda/10$ environ). En observant les équations des lignes, on constate qu'elles peuvent être traduites en un circuit équivalent donné sur la figure 2.10. La résolution peut être réalisée par une approche de type différences finies (présentée plus loin dans le mémoire pour un problème tridimensionnel). Les courants et tensions induits au niveau de chaque section en fonction du temps constituent les inconnues du problème.

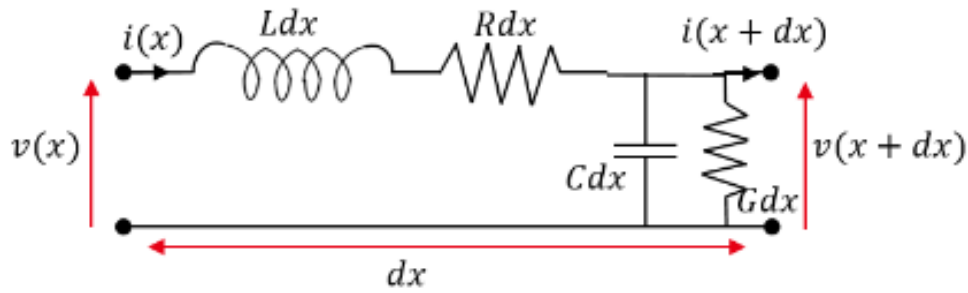


Figure 2.10: Circuit électrique équivalent à un tronçon de ligne de transmission, montrant les paramètres linéiques R , L , C et G .

- R représente la résistance série des conducteurs, en Ω/m .
- L représente l’inductance série de la ligne, en H/m ,
- G représente les pertes dans le diélectrique qui sépare les conducteurs, en S/m ,
- C représente la capacité linéique, en F/m .

En appliquant les lois de Kirchhoff à ce circuit, on peut déterminer la tension et le courant, en considérant que les variations dans le temps de $v(x)$ et $v(x + dx)$ sont égales :

$$v(x + dx, t) - v(x, t) = -Rdx i(x, t) - Ldx \frac{\partial i(x, t)}{\partial t} \quad (2.2)$$

$$i(x + dx, t) - i(x, t) = -Gdx v(x, t) - Cdx \frac{\partial v(x, t)}{\partial t} \quad (2.3)$$

Dans la ligne de longueur dx , i et v représente respectivement le courant et la tension à l’instant t . En divisant chaque élément des équations (1.1) et (1.2) par dx et en faisant tendre Δx vers zéro, nous obtiendrons les équations suivantes .

$$\frac{\partial v(x, t)}{\partial x} = -Ri(x, t) - L \frac{\partial i(x, t)}{\partial t} \quad (2.4)$$

$$\frac{\partial i(x, t)}{\partial x} = -Gv(x, t) - C \frac{\partial v(x, t)}{\partial t} \quad (2.5)$$

2.7.2 Solutions générales en régime sinusoïdal

En dérivant ces équations de nouveau par rapport à x puis à t , nous obtiendrons l’équation des télégraphistes. En régime harmonique, on suppose que l’excitation est une onde sinusoïdale de pulsation ω . La tension et le courant à chaque

position sur la ligne peuvent être formulés comme suit :

$$v(x, \omega, t) = V(x, \omega)e^{j\omega t} \quad (2.6)$$

$$i(x, \omega, t) = I(x, \omega)e^{j\omega t} \quad (2.7)$$

v et i représentent la tension et le courant instantané en tout point de la ligne, V et I représentent les variables de la tension et du courant dans le domaine spectral. En remplaçant $v(x, t)$ et $i(x, t)$ dans les équations (1.6) et (1.7), on obtient:

$$\frac{\partial^2 V(x, t)}{\partial x^2} = \gamma^2 V(x, t) \quad (2.8)$$

$$\frac{\partial^2 I(x, t)}{\partial x^2} = \gamma^2 I(x, t) \quad (2.9)$$

En définissant le coefficient de propagation linéaire comme suit :

$$\gamma = (G + j\omega C)(R + j\omega L) = \alpha + j\beta \quad (2.10)$$

- γ est appelée la constante de propagation complexe.
- α représente la constante linéaire (Neper/m)
- β représente la constante de phase (radian/m).

En résolvant ces équations dans le domaine fréquentiel, on obtient les équations suivantes:

$$V^+(x, t) = V^+ e^{-\gamma x} + V^- e^{\gamma x} = Z_c(\omega)I(x, t) \quad (2.11)$$

$$I(x, t) = I^+ e^{-\gamma x} + I^- e^{\gamma x} \quad (2.12)$$

où le terme $e^{-\gamma x}$ représente la propagation de l’onde dans la direction $+x$, et le terme $e^{\gamma x}$ représente la propagation de l’onde dans la direction $-x$. Les ondes de tension $V(z, t)$ et de courant $I(z, t)$ sont associées par une impédance caractéristique Z_c de la ligne, définie comme :

$$Z_c = \sqrt{\frac{R + j\omega L}{G + j\omega C}} \quad (2.13)$$

Cas particulier de la ligne sans perte

Dans le cas où la ligne ne présente aucune perte, la constante de propagation peut être exprimée comme suit:

$$\gamma = j\omega\sqrt{LC} = j\beta \quad (2.14)$$

L’impédance caractéristique s’exprime comme suit:

$$Z_c = \sqrt{\frac{L}{C}} \quad (2.15)$$

2.7.3 Coefficient de réflexion

La forme de la tension (eq 1.11) montre que l’onde est décomposée en deux ondes : une onde progressive et une onde rétrograde, le rapport entre ces ondes progressive et réfléchie représente le coefficient de réflexion. Dans cette section, nous détaillons le coefficient de réflexion. Chaque onde réfléchie résulte d’une discontinuité (comme des jonctions, des défauts, etc.). Considérons une ligne de transmission de longueur l avec une impédance caractéristique Z_c . Cette ligne est alimentée par un générateur de tension à l’extrémité gauche, comme illustré sur la figure 2.11, et est terminée à l’extrémité droite par une impédance notée Z_t . Le coefficient de réflexion, qui est le rapport entre l’onde incidente et l’onde réfléchie, est désigné par Γ_t .

$$\Gamma_t = \frac{V^-}{V^+} = \frac{Z_t - Z_g}{Z_t + Z_g} \quad (2.16)$$

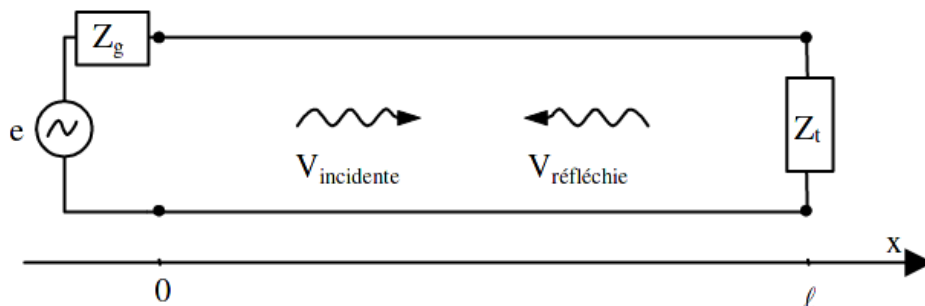


Figure 2.11: Modèle d’une ligne de transmission de longueur L et d’impédance caractéristique Z_t .

La réflectométrie se base principalement sur l’analyse des signaux réfléchis :
Court-circuit (Défaut franc):

Si une ligne de transmission est court-circuitée à son extrémité, l’impédance de charge Z_t est de 0 ohms, $\Gamma = -1$.

Circuit ouvert (Défaut franc):

Si une ligne est en circuit ouvert à son extrémité, l’impédance de charge Z_t tend vers l’infini $\Gamma = 1$.

Adaptation d’impédance:

Si la charge est parfaitement adaptée à la ligne, c’est-à-dire que $Z_t = Z_g$, il n’y a pas d’onde réfléchie $\Gamma = 0$.

Défauts non francs:

Ce type de défauts se caractérise par une faible amplitude des échos, contrairement aux défauts francs tels que les courts-circuits ou les circuits ouverts. Ces caractéristiques les rendent plus difficiles à détecter. On distingue principalement deux types d’ondes : les ondes réfléchies, renvoyées vers la source ou le point d’injection, et celles qui continuent à se propager le long de la ligne de transmission. Dans ce contexte, le coefficient de transmission est défini comme le rapport entre l’onde transmise à la charge et l’onde incidente.

$$T_t = \frac{V(x, t)}{V^+ e^{-\gamma x}} \quad (2.17)$$

A partir de l’équation (2.17), on peut déduire que:

$$T_t = 1 + \Gamma_t \quad (2.18)$$

En mesurant le temps écoulé entre l’émission de l’onde initiale et la réception de l’onde réfléchie, on peut déterminer la position du défaut le long de la ligne à partir de l’équation suivante :

$$d = \frac{\tau v_p}{2} \quad (2.19)$$

avec

- τ correspond au temps nécessaire pour parcourir la ligne sur un aller-retour
- v_p est la vitesse de propagation dans la ligne.

La vitesse de propagation dépend principalement des propriétés électriques du matériau utilisé ainsi que du milieu de propagation. Dans un cas idéal, où la ligne de transmission ne présente aucune perte, la vitesse est définie par :

$$v_p = \frac{1}{\sqrt{LC}} \quad (2.20)$$

avec

- L est l’inductance par unité de longueur (en henry par mètre, H/m).
- C est la capacité par unité de longueur (en farads par mètre, F/m).

2.7.4 Modélisation de défauts

Nous avons vu comment réaliser la modélisation d’un tronçon de ligne de longueur dx . La modélisation de la ligne complète sous la forme d’un schéma électrique est réalisée en disposant ces tronçons bout à bout. Ainsi si un défaut non franc est présent le long de la ligne, il va se présenter comme la variation d’un des quatre paramètres linéiques du tronçon sur lequel il se trouve. Ainsi par exemple, on peut considérer une variation de capacité locale. De ce fait, nous aurons un changement local de l’impédance qui impactera directement les coefficients de transmission et de réflexion à l’aide des formules données plus haut.

Les défauts au sein des lignes de transmission peuvent être catégorisés en trois types principaux : les défauts répartis (ou distribués), les défauts en série et les défauts en parallèle. Une manière efficace de représenter ces défauts est d’intégrer une variante altérée du réseau à 2 ports illustré sur la Figure 2.12.

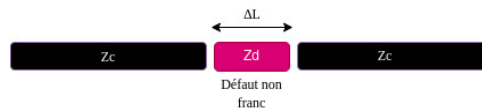


Figure 2.12: Modélisation du défaut distribué

Dans cette variante, les paramètres L , C , R et G et sont ajustés en fonction de l’intensité du défaut. De plus, le défaut est défini par une longueur L_0 et une impédance

$$Z_{df} = Z_c(1 \pm \Delta Z_c) \quad (2.21)$$

où $-Z_c < \pm \Delta Z_c < Z_c$ comme illustré sur la Figure 2.12.

On peut constater que l’introduction d’un défaut dans une branche influence directement son impédance. Pour une branche défectueuse, cette modification se traduit par les équations :

$$\Gamma_{in} = \frac{Z_d - Z_c}{Z_d + Z_c} \quad (2.22)$$

et

$$\Gamma_{out} = \frac{Z_c - Z_d}{Z_c + Z_d} = -\Gamma_{in}. \quad (2.23)$$

Lorsqu’une onde incidente arrive à l’extrémité gauche d’un défaut, la tension réfléchie est caractérisée par Γ_{in} . Par ailleurs, la tension qui traverse le défaut est représentée par $1 + \Gamma_{in}$. Il y a une multitude de réflexions entre les côtés gauche et droit du défaut. Chaque interaction a un décalage temporel de $e^{-\gamma \times l}$, où l est la longueur sur laquelle s’étend le défaut, et γ est la constante de propagation qui décrit la manière dont l’onde se propage à l’intérieur du défaut. Le coefficient de réflexion global du défaut est décrit dans la référence [64].

$$\Gamma_{df} = \Gamma_{in} \times \frac{1 - \exp^{-2\gamma l}}{1 - \Gamma_{in}^2 \exp^{-2\gamma l}} \quad (2.24)$$

2.8 Méthodes basées sur la réflectométrie

Les méthodes de détection de défauts dans les câbles utilisent la réflectométrie. Il existe deux grandes classes : l’analyse dans le domaine du temps, et dans le domaine fréquentiel.

2.8.1 Réflectométrie dans le domaine temporel

Dans le domaine temporel, on peut distinguer trois catégories d’approches : la réflectométrie classique utilisée dans le cadre de cette thèse, la S/SSTDR (réflectométrie par séquence directe et par étalement de spectre) et la MCTDR (Multi-carrier Time Domain Reflectometry).

2.9 Réflectométrie classique

La méthode de réflectométrie classique repose sur l’injection d’une impulsion gaussienne dans la ligne à tester. Cette impulsion agit comme un signal sonde, et les réflexions causées par les défauts de la ligne, revenant à l’entrée de la ligne, sont analysées pour obtenir des informations sur leur nature et leur emplacement. La forme temporelle de cette impulsion gaussienne est décrite mathématiquement par l’équation :

$$G(t) = A \cdot e^{-\frac{1}{2} \left(\frac{t}{a} \right)^2} \quad (2.25)$$

où :

- A représente l’amplitude crête de l’impulsion gaussienne.

– a est le facteur d’étalement, qui détermine la largeur de l’impulsion. Plus la valeur de a est grande, plus l’impulsion est étalée dans le temps et vice-versa. La Figure 2.13 présente une représentation schématique du dispositif de réflectométrie classique exploité au sein de cette thèse. Au cœur de cette illustration se trouve le coupleur, élément central en réflectométrie. Sa mission principale est de guider l’impulsion créée vers la ligne de test, tout en s’assurant que l’impulsion qui est réfléchiée est correctement dirigée vers l’unité d’acquisition. Le générateur d’impulsion est la source d’une impulsion gaussienne. Une fois générée, cette impulsion est injectée et se propage le long de la ligne de test. Si un défaut est présent sur cette ligne, une portion de l’impulsion est réfléchiée. L’unité d’acquisition, finalement, permet de mesurer le retard avant le retour de l’impulsion réfléchiée ainsi que son amplitude.

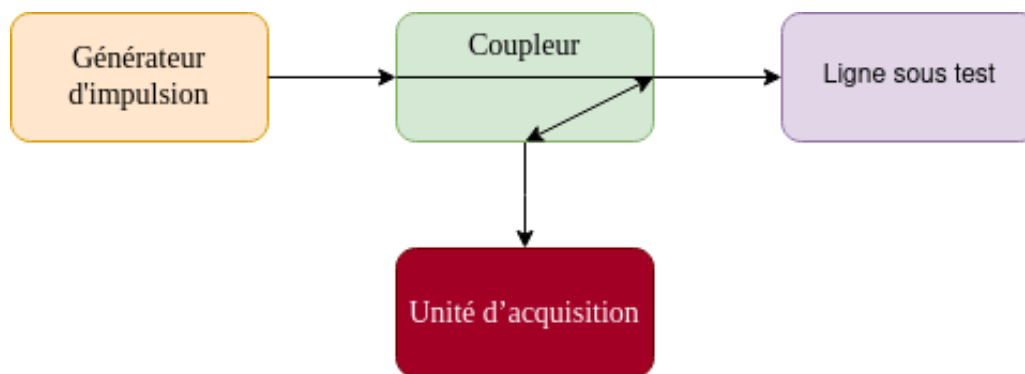


Figure 2.13: Principe de réflectométrie classique

2.9.1 STDR/SSTDR

La méthode STDR (Spread Time Domain Reflectometry) et la méthode SSTDR (Sine Sweep Time Domain Reflectometry) [65] [66] sont utilisées pour détecter et localiser les défauts dans les câbles et les réseaux de transmission. Nous expliquons dans cette section le principe de fonctionnement de chaque approche ainsi que leurs avantages et inconvénients.

Méthode STDR/SSTDR

La principale différence entre STDR et SSTDR réside dans le type de signal PN utilisé. Dans STDR, on injecte un pseudo-bruit (PN) simple dans le câble, tandis que dans SSTDR, on injecte une séquence de PN modulée par une onde sinusoïdale. Cette modulation ajoute une composante sinusoïdale au signal, ce qui

peut faciliter la détection de certains types de défauts ou variations d’impédance. En revanche, STDR offre une approche plus simple et directe en utilisant uniquement le pseudo-bruit. Les deux méthodes visent néanmoins à atteindre le même objectif, qui est de détecter et localiser les défauts dans les câbles.

MCTDR

La Réflectométrie Multicarrier en Domaine Temporel (MCTDR) [67], a pour objectif de simplifier la mise en œuvre de diagnostics en temps réel, c’est-à-dire pendant les phases de fonctionnement du système en question. En mode de fonctionnement, il est important que le signal de test n’interfère pas le signal utile. Cela peut être nécessaire pour des raisons de compatibilité électromagnétique ou lorsque les fréquences en question seraient la cause de trop de distorsion dans les mesures. Pour résoudre ce problème, une méthode appelée "Noise Domain Reflectometry" utilise les signaux déjà présents ainsi que du bruit comme stimulation, sans ajouter de nouveaux signaux. Cependant, il est rare que les signaux dans les câbles à tester ressemblent à un bruit aléatoire.

La méthode MCTDR (Modulated Continuous Time Domain Reflectometry) fonctionne en ajustant les propriétés spectrales du signal injecté. Cela permet d’éviter les interférences entre le signal de test et les signaux utiles circulant dans le câble en cours d’utilisation. Le signal de test pour un système donné consiste en la combinaison ajustée en amplitude et en phase de différentes sinusoïdes.

2.9.2 Réflectométrie dans le domaine fréquentiel

Dans le domaine fréquentiel, le but de cette technique est d’utiliser un ensemble de signaux sinusoïdaux à fréquence échelonnée dans le temps, connus sous le nom de signaux chirps, pour localiser les défauts. La détection des défauts résulte de l’analyse soit du signal réfléchi, soit du signal stationnaire engendré par l’interférence entre le signal injecté et le signal réfléchi. Pour ce faire, on injecte un signal de type chirp, dont la représentation est donnée par l’expression suivante et illustré sur la figure 2.14:

$$x_i(t) = A \sin(\theta(t) + \varphi), \text{ avec: } \theta(t) = \int_0^t 2\pi f(u) du \quad (2.26)$$

$$f(t) = f_{min} + \frac{f_{max} - f_{min}}{t_{max}}, \text{ avec: } 0 \leq t \leq t_{max} \quad (2.27)$$

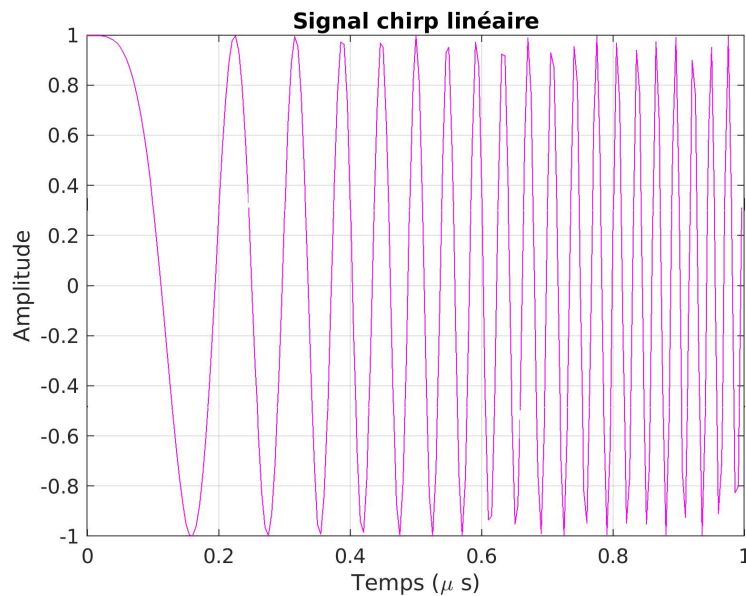


Figure 2.14: Exemple de signal chirp linéaire avec son spectre fréquentiel couvrant une bande passant de 40MHz.

- La phase et la fréquence instantanée $\mathcal{O}(t)$ et $f(t)$ représente les paramètres clés qui varient dans la plage de fréquence allant de f_{min} et f_{max}

Pour identifier la présence des discontinuités et les défauts dans le réseau de câbles, les techniques FDR (Frequency Domain Reflectometry) [68] permettent une évaluation directe de la réponse du réseau dans la plage de fréquences. Cette évaluation se fait en analysant la fréquence, la phase ou l’amplitude afin de détecter les variations d’impédance.

Modulation de fréquence par ondes continues (FMCW)

La Réflectométrie à Onde Continue Modulée en Fréquence (FMCW) [51] est une méthode de détection de défauts utilisant un Oscillateur Commandé par Tension (VCO) ou un autre générateur d’ondes sinusoïdales. Son principe repose sur l’injection périodique d’une onde sinusoïdale à haute fréquence.

Lorsque cette onde sinusoïdale rencontre une discontinuité dans le câble, elle se réfléchit et revient au point d’injection. Un coupleur directionnel est alors utilisé pour distinguer l’onde incidente de l’onde réfléchie.

Le but de cette méthode est de mesurer le décalage de fréquence ΔF entre le signal incident et le signal réfléchi, retardé temporellement de Δt . Cette mesure permet de détecter et localiser les défauts tels que les discontinuités, les défauts d’impédance ou les changements de caractéristiques dans le câble.

La relation entre ΔF et Δt est donnée par une équation spécifique qui permet d’analyser le comportement du signal réfléchi et ainsi identifier la présence et les caractéristiques des défauts dans le câble.

$$\Delta F = \frac{\Delta t}{T} (f_{max} - f_{min}) \quad (2.28)$$

L’inconvénient de la méthode FMCW (Réflectométrie à Onde Continue Modulée en Fréquence) pour la détection des défauts non francs est qu’elle est principalement efficace pour des coefficients de réflexion d’amplitude significative. Les défauts non francs, tels que les discontinuités de faible amplitude ou les défauts d’impédance subtils, peuvent ne pas provoquer une réflexion suffisamment forte pour être détectés de manière fiable par cette méthode.

Étant donné que la méthode FMCW repose sur la mesure du décalage de fréquence entre le signal incident et le signal réfléchi, les défauts non francs qui ne génèrent qu’une faible réflexion peuvent être difficiles à distinguer du bruit de fond ou des variations de signal causées par d’autres facteurs.

Réflectométrie à Onde Stationnaire (SWR)

Le but de la méthode SWR (Standing-Wave Reflectometry) [69] est de localiser et identifier les défauts ou charges sur un câble en utilisant une onde sinusoïdale à haute fréquence. La méthode SWR consiste à injecter un signal sinusoïdal à haute fréquence dans le câble. Contrairement aux autres méthodes de réflectométrie, la méthode SWR ne sépare pas l’onde injectée de l’onde réfléchie, ce qui entraîne la formation d’une onde stationnaire (standing wave) sur le câble. Cette onde stationnaire est créée par la superposition de l’onde incidente et de l’onde réfléchie, produisant des minima et de maxima d’amplitude. Le processus pour déterminer la localisation et le type de charge sur laquelle la réflexion s’est produite consiste à analyser le comportement d’amplitude de l’onde stationnaire aux points de minimum. Ces points de minimum d’amplitude fournissent des informations sur la présence et les caractéristiques du défaut ou de la charge.

L’inconvénient majeur de la méthode SWR réside dans sa sensibilité au bruit. Étant donné que cette méthode se base sur l’analyse d’amplitude aux points de minimum, toute interférence ou bruit présent dans le câble peut affecter les mesures et entraîner des fausses détections ou des résultats imprécis. Une autre procédure peut être utilisée pour localiser un défaut en envoyant un ensemble de signaux sinusoïdaux de fréquence progressive et en observant les points

de minimum et de maximum, ce qui peut potentiellement réduire l’effet du bruit. Cependant, malgré cette approche alternative, la sensibilité au bruit reste un défi majeur de la méthode SWR.

Détection de Phase en Réflectométrie à Domaine Fréquentiel (PDFDR)

Principe de la méthode PDFDR:

La méthode PDFDR [69], [70] est basée sur l’évaluation de la phase du signal réfléchi par rapport au signal incident dans le domaine fréquentiel pour détecter des discontinuités ou des défauts dans un câble, lesquels provoquent des réflexions et changent la phase du signal réfléchi. En appliquant cette méthode, un signal électromagnétique est envoyé dans le câble et la différence de phase entre le signal réfléchi et le signal incident est analysée pour déterminer la position des défauts. Ceci permet d’obtenir une représentation du profil d’impédance du câble et de mettre en évidence les variations d’impédance causées par les défauts. Cependant, la méthode PDFDR présente une limitation significative: elle est principalement efficace pour détecter les défauts francs, qui entraînent des réflexions importantes et des changements de phase significatifs. Les défauts non francs, tels que les variations d’impédance subtiles ou de faible amplitude, peuvent ne pas induire des changements de phase suffisamment distincts, rendant ainsi leur détection moins fiable.

2.10 Conclusion

Ce chapitre est principalement consacré à l’exploration des aspects et parties théoriques de notre étude : la propagation le long de la ligne de transmission et la méthode d’analyse par réflectométrie. Ces concepts fondamentaux formeront le socle sur lequel reposera la génération de notre base de données. Cette dernière sera par la suite exploitée dans le but de localiser et d’étudier les défauts, en particulier les défauts non francs, en utilisant des techniques avancées de réseaux de neurones tels que les MLP (Perceptrons multicouches) et les CNN (Réseaux Neuronaux Convolutifs).

Tout au long de ce chapitre, nous avons fait un bilan des approches d’analyse de défaut par les mécanismes de réflectométrie, en explorant son principe et en se basant sur la théorie des lignes de transmission. Nous avons également

évoqué les divers types de défauts, qu’ils soient francs ou non, et comment ils se manifestent.

De plus, ce chapitre a également été une occasion d’examiner et de distinguer les avantages et inconvénients des différentes méthodes de diagnostic proposées, que ce soit le diagnostic en ligne (tel que la réflectométrie par étalement de spectre et la réflectométrie multicarriers).

À travers cette exploration théorique, nous établissons donc les principes nécessaires à l’avancée de notre étude pratique, qui s’articulera autour de l’application des techniques d’intelligence artificielle pour le diagnostic et la localisation des défauts dans les réseaux filaires.

3

Diagnostic de l'état d'un réseau filaire complexe

Sommaire

3.1	Introduction	92
3.2	Topologie de câbles complexe	92
3.3	Description de la base de données	95
3.4	Mise en oeuvre des réseaux de neurones classiques pour la détection des défauts non francs	98
3.5	Mise en oeuvre des réseaux de neurones convolutifs pour la détection des défauts non francs	107
3.6	Sélection des hyper-paramètres optimaux	107
3.7	Synthèse	111

3.1 Introduction

Dans cette section, nous allons présenter une méthodologie permettant l'application de réseaux de neurones (RN) dans le domaine du diagnostic d'un réseau de câbles complexe. La détection des défauts le long des câbles est réalisée par réflectométrie dans le domaine temporel (TDR pour Time Domain Reflectometry). Dans ce cadre. Le réseau qui sera étudié, comme le montre la Figure 3.1, est composé de câbles coaxiaux. Dans ce contexte, nous nous concentrerons uniquement sur l'analyse d'un type particulier de défauts qui difficile à détecter : les défauts non francs.

3.2 Topologie de câbles complexe

Notre réseau est constitué de câbles coaxiaux d'impédance caractéristique $Z_c = 50 \Omega$, diagnostiqués par deux réflectomètres $R1$ et $R2$. Il est composé de 5 câbles interconnectés. Les longueurs de chaque brin sont données sur la Figure 3.1. Le réseau est affecté par deux défauts non francs choisis de façon aléatoire aussi bien pour le choix de la branche que celui de la position sur la branche.

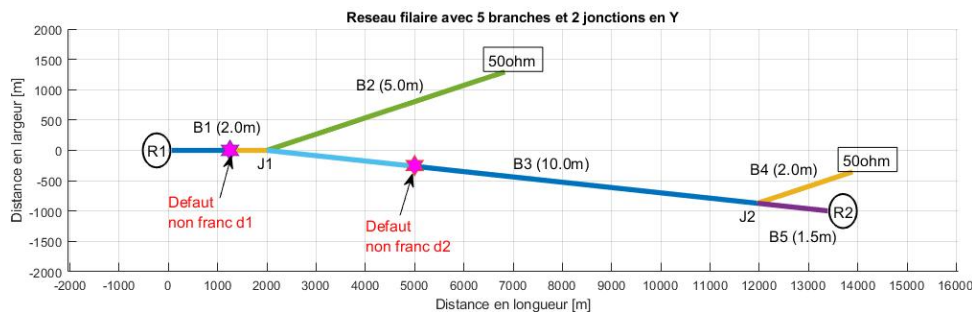


Figure 3.1: Réseau bifilaire à 5 branches et deux défauts non francs sur la branche B1 et la branche B3

Dans le cas présent, le défaut $d1$ sur la branche B1 est situé à une distance de 1,25 m du réflectomètre R1 sur la branche B1, tandis que le défaut $d2$ est situé sur la branche B3 à une distance de 3 m du même réflectomètre R1.

Le protocole de mesure hypothétique est le suivant : dans un premier temps, un signal serait injecté sur le réflectomètre R1 (R2 absent dans ce cas), d'où on déduirait le réflectogramme correspondant au niveau de R1. Pour cette simulation, l'extrémité de la ligne où se trouverait R2 (R1 également absent) serait adaptée. Dans un second temps, un signal serait injecté au niveau hypothétique de R2 et le

réflectogramme serait visualisé à ce niveau, l'extrémité où se trouverait R1 étant adaptée. Cela réaliserait ainsi deux mesures successives. Le choix d'utiliser deux réflectomètres plutôt qu'un est dicté par le fait que dans un réseau complexe, on peut se retrouver avec des ambiguïtés. En effet, prenons l'exemple d'un défaut mesuré à une distance de 3m de R1. Alors, à priori, ce défaut peut très bien se trouver sur la branche B2 ou sur la branche B3. Le protocole choisi lèverait ce type d'ambiguïté.

La Figure 3.2 illustre la réponse temporelle du réseau sain et du réseau défectueux sur le réflectomètre R1 et la Figure 3.3 illustre la réponse temporelle du réseau sain et du réseau défectueux sur le réflectomètre R2.

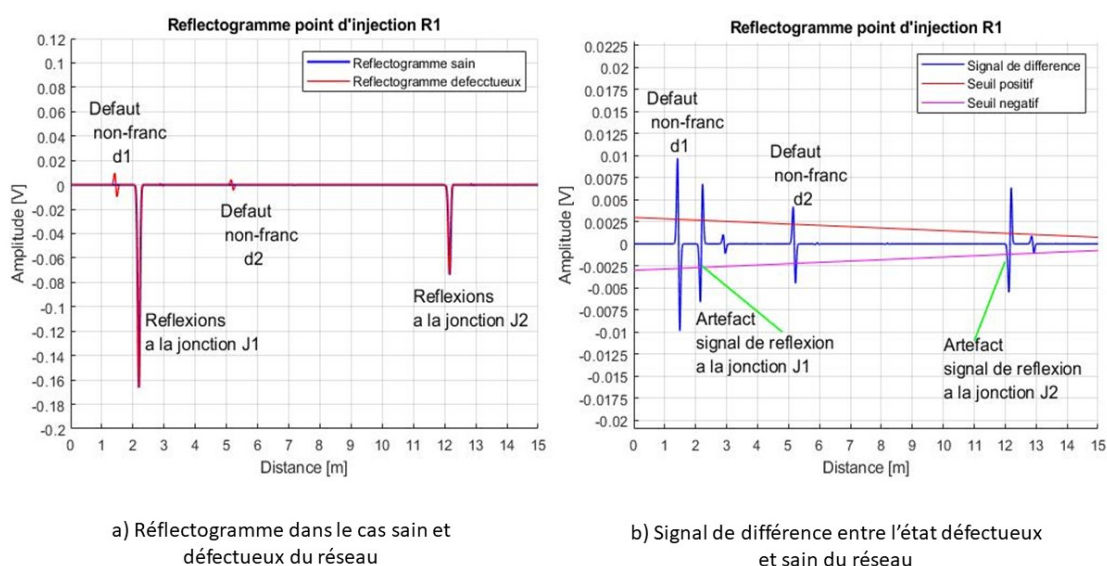


Figure 3.2: Réflectogramme au niveau de R1

Un signal gaussien d'amplitude 1V et de fréquence centrale de 1GHz a été utilisé pour exciter les câbles aux points d'injection des réflectomètres R1 et R2. Dans une mesure, un coupleur en forme de T est inséré aux points d'injection R1 et R2 empêchant les excitations à la source d'apparaître à l'entrée du récepteur du réflectomètre. Ainsi, on ne voit pas les signaux à la source sur les réflectogrammes R1 et R2 captés par les récepteurs VNA (lorsqu'on réalise une mesure plutôt qu'une simulation). En revanche, les signaux négatifs de grande amplitude sur le réflectogramme R1 dans la figure 1.2 a) correspondent à la jonction J1 à une distance de 2m et J2 à une distance de 12m, respectivement. De même, les signaux négatifs de grande amplitude sur le réflectogramme R2 sur la figure 1.3

a) correspondent à la jonction J2 à une distance de 1.5m et J1 a une distance de 11.5m, respectivement.

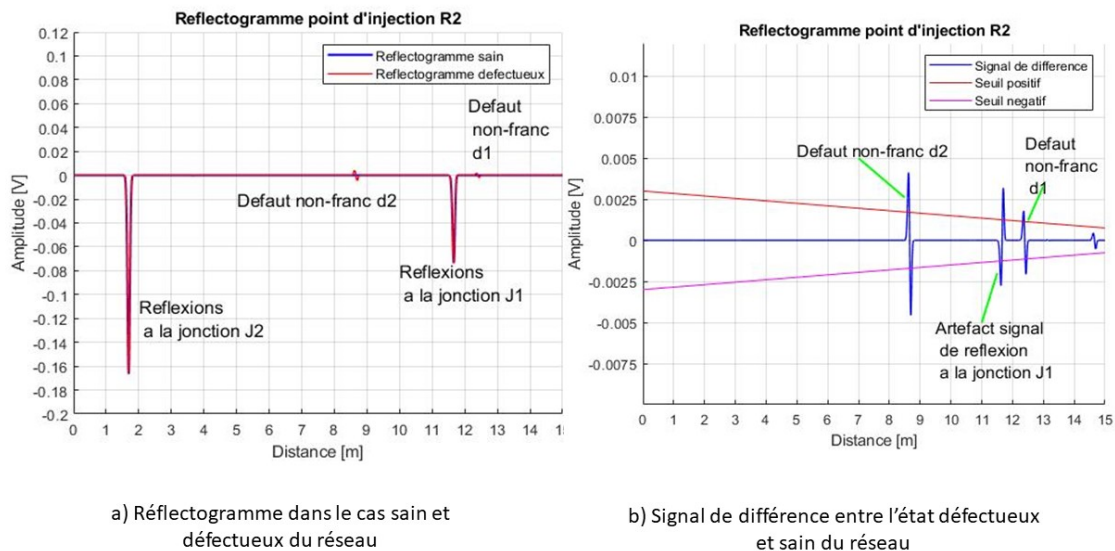


Figure 3.3: Réflectogramme au niveau de R2

Par exemple, deux défauts non francs sont détectés par la réflectométrie R2. Le premier est situé à une distance de 1,5 m près de la source d'excitation et le second à une distance de 12,5 m.

3.2.1 Ambiguïté de localisation des défauts

Le phénomène d'atténuation empêche un seul point d'injection de couvrir efficacement l'ensemble d'un réseau complexe. Ce problème est encore compliqué par les nombreuses branches du réseau, qui créent une ambiguïté lors de la localisation des défauts. Une solution possible à ces problèmes est la réflectométrie distribuée, qui améliore la précision et la qualité du diagnostic en répartissant les mesures de diagnostic sur plusieurs points du réseau. Diverses solutions pour analyser à la fois les lignes de transmission simples et les topologies complexes de réseaux câblés sont disponibles dans la littérature.

Par exemple, sur la figure 1.2 le défaut non franc d_2 apparaît à une distance de 5m sur le réflectogramme R_1 par rapport à la position du réflectomètre R_1 . Pour le capteur R_1 , le défaut d_2 peut se trouver sur la branche B_2 ou la branche B_3 , après

la jonction J_1 mais avant la jonction J_2 . Dans ce cas on a donc deux branches, B_2 et B_3 , ambiguës pour le défaut d_2 et le réflectomètre R_1 ne peut pas éliminer cette ambiguïté concernant la localisation du défaut d_2 . Avec un deuxième réflectomètre R_2 installé à l'extrémité de la branche B_5 on produit le réflectogramme R_2 où, comme illustré sur la figure 1.3, le défaut franc d_2 se trouve entre les jonctions J_2 et J_1 . Ainsi, après le post-traitement des deux réflectogrammes R_1 et R_2 on déduit le fait que le défaut d_2 est localisé sur la branche B_3 , à 5m de distance par rapport au point d'injection R_1 de la branche B_1 . On peut ainsi constater que la fusion des informations des réflectomètres distribués aide à surmonter les difficultés de localisation avec précision des défauts non francs dues aux solutions multiples et qui sont résolues en augmentant le nombre de capteurs.

3.3 Description de la base de données

Nous avons généré une base de données pour la création, le test et la validation d'un modèle de diagnostic à base de réseau neuronal. Notre base de données comprend 37 colonnes, parmi lesquelles 32 colonnes représentent les données d'entrée de notre réseau neuronal. Ces données d'entrée illustrent les signaux réfléchis capturés par deux réflectomètres, R_1 et R_2 . Ils consistent en l'amplitude du signal réfléchi liée aux défauts ainsi qu'en leurs positions mesurées. Quant aux 5 colonnes restantes, elles correspondent aux 5 branches sur lesquelles est précisé où le défaut a été localisé.

La figure 3.1 illustre le réseau à 5 branches utilisé pour la génération de la base de données. Le réseau est constitué des branches B_1 (2.0m), B_2 (5.0m), B_3 (10.0m), B_4 (1.5m) et B_5 (2.0m) qui sont connectées via deux jonctions : J_1 étant la jonction des branches B_1 , B_2 , et B_3 et J_2 étant la jonction des branches B_3 , B_4 et B_5 . Un réflectomètre R_1 est placé au début de la branche B_1 et un deuxième réflectomètre R_2 est placé à l'extrémité de la branche B_5 . À chaque simulation, nous générons un défaut sur une branche choisie aléatoirement, dont la position est également choisie de manière aléatoire. La Figure 1.4 illustre la génération d'un défaut non franc, positionné aléatoirement sur la branche B_3 .

Les réflectogrammes captés par R_1 et R_2 permettent de détecter le défaut non franc et de le localiser sur la branche B_3 . Ainsi, l'ambiguïté de localisation uniquement avec le réflectogramme de R_1 sur la branche B_2 ou la branche B_3 est résolue grâce à la localisation indiquée par le réflectogramme R_2 .

Initialement, nous avons généré 2000 échantillons d'observation. Cependant,

pour éviter un surajustement de notre modèle, nous avons réduit les données à 1000 échantillons.

La figure 3.4 montre le signal réfléchi par le toron défectueux sur la branche B3 et détecté par R2, ainsi que l'artefact à la jonction J1, résultant de la différence entre le réflectogramme sain R2 et le réflectogramme R2 avec un défaut non franc.

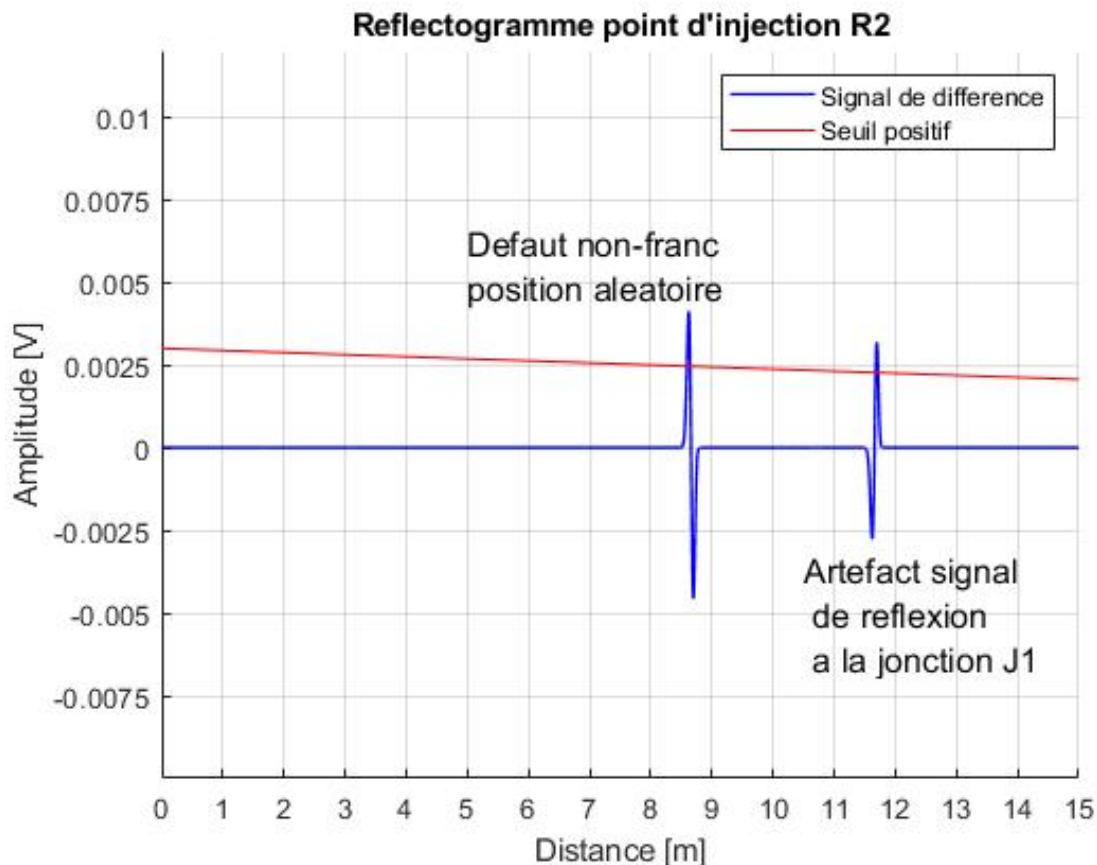


Figure 3.4: Les deux pics du défaut non franc, positionné aléatoirement et détecté par R2 sur la branche B3.

Le défaut non franc possède deux pics qui sont déterminés par la différence entre la partie du toron sain de la branche et la petite portion de 2cm du toron défectueux : un pic positif à l'entrée du signal émis par R2 dans le toron défectueux de 2cm et un pic négatif à la sortie, quand le signal reprend la propagation sur la portion saine du câble. Le pic positif du début de défaut montre que l'impédance de la portion de câble défectueux a une impédance caractéristique est plus élevée que Z_c , l'impédance caractéristique du câble sain.

Le changement de vitesse de propagation dans le petit toron qui modélise le défaut non franc fait qu'à l'endroit de la jonction J1, le réflectogramme du réseau

défectueux a un déphasage par rapport au réflectogramme du réseau sain : par conséquent, les signaux réfléchis à l'endroit de la jonction $J1$ ne sont plus en phase et leur différence fait apparaître l'artéfact qu'on voit après le signal généré par le défaut non franc situé avant la jonction $J1$. Ces artefacts créent des problèmes dans la détection des signaux de défauts et leur localisation.

Une réflectogramme similaire à celle de la figure 1.5 est également obtenue par $R1$. Après avoir obtenu ces réflectogrammes $R1$ et $R2$ par la soustraction des réflectogrammes saines de celui défectueux, nous détectons les pics des signaux réfléchis par le défaut et nous les stockons dans la base de données par pair : position-amplitude positive du pic.

Afin de présenter la détection des pics des signaux réfléchis par des défauts non francs, nous montrons dans la figure 3.5 la forme du signal réfléchi par le toron défectueux :

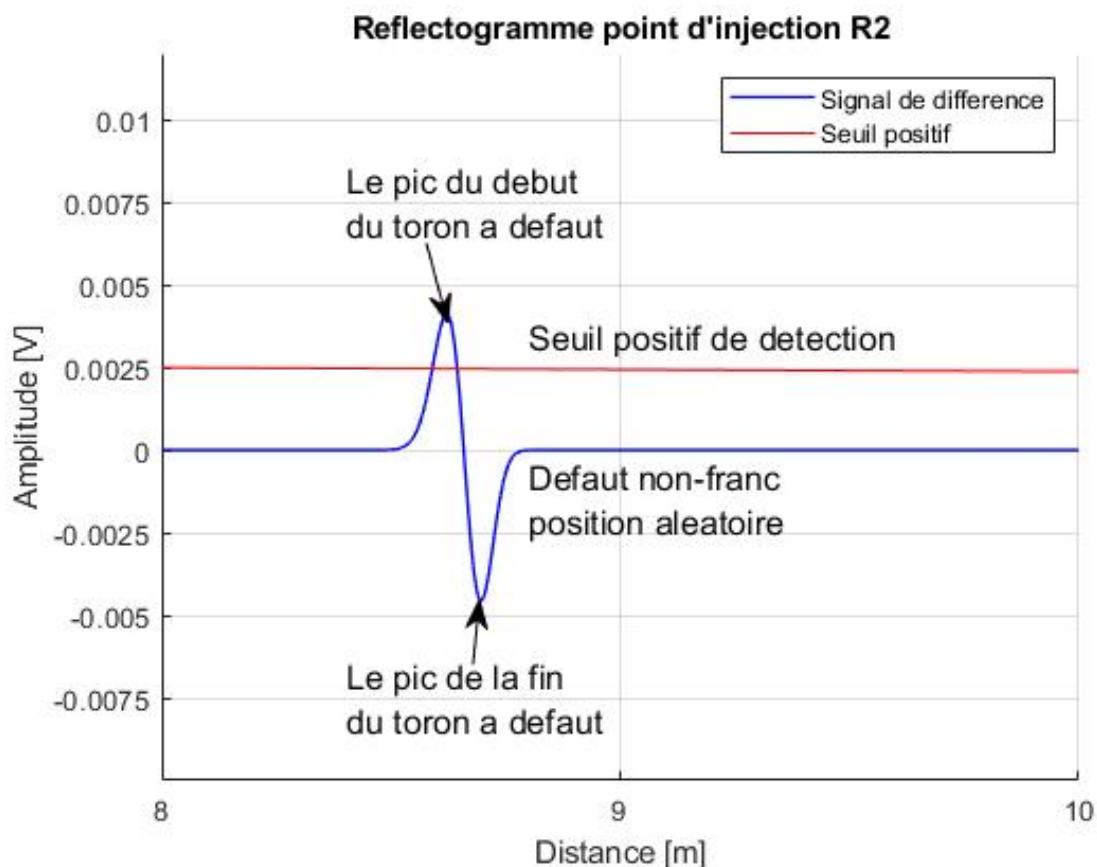


Figure 3.5: La forme d'onde du signal réfléchi par le défaut non franc : a) à gauche, le pic positif à l'entrée du toron défectueux; b) à droite, le pic négatif à la sortie du toron défectueux

Dans la base de données, nous détectons et stockons les valeurs positives (ou

absolues) des deux pics qui apparaissent aux endroits où le câble sain est en contact avec la portion du câble affectée par le défaut non franc. Pour la détection des valeurs absolues des deux pics, nous comparons ces valeurs avec un seuil positif, dont la valeur est choisie en fonction de paramètres tels que :

- La valeur minimale ΔZ_c à détecter pour la zone à défaut non franc.
- La longueur des câbles.
- L'atténuation des câbles.
- Le bruit maximal toléré.

La détection des pics commence par la comparaison des valeurs absolues des échantillons avec le seuil positif ; dès que les échantillons dépassent la valeur du seuil, nous comparons l'échantillon actuel avec le précédent et conservons le plus grand des deux. Lorsque le plus grand est le prédécesseur, celui-ci est stocké comme le pic détecté.

Après la détection des pics, nous les stockons dans un fichier de la base de données sous la forme de paires (position, amplitude), deux paires pour le réflectogramme $R1$ et deux paires pour le réflectogramme $R2$. Sur les cinq colonnes de sortie, nous inscrivons un 1 dans la colonne de la branche où le défaut a été généré et conservons les autres colonnes des branches sans défaut à 0.

3.4 Mise en oeuvre des réseaux de neurones classiques pour la détection des défauts non francs

3.4.1 Modèle de régression

La Figure 3.6 présente un schéma détaillant la méthode de diagnostic d'un réseau filaire sophistiqué en utilisant une approche basée sur les réseaux de neurones (RN). Les signaux TDR (trace de réflexion) du réseau filaire servent de données d'entrée au RN, tandis que les informations de sortie correspondent à la position du défaut.

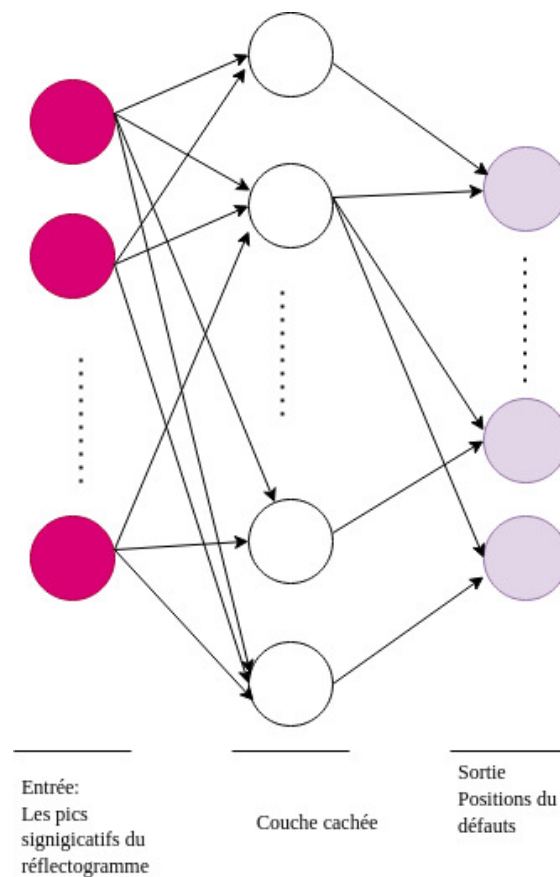


Figure 3.6: Schéma RN

Avant de procéder à l'entraînement de notre modèle de régression, une première étape importante consiste à préparer l'ensemble des données. Cette phase est indispensable car un modèle bien entraîné dépend des données sur lesquelles il est construit. Dans ce contexte, nous utilisons des dataframes qui permettent une manipulation structurée de l'ensemble de données. Afin de garantir un bon entraînement de notre modèle, nous allons normaliser toutes les caractéristiques pour faciliter la convergence du modèle. Ce processus permettra aux optimiseurs, comme la descente de gradient, de converger plus rapidement lorsque tous les attributs se trouvent sur des échelles similaires. Pour ce faire, nous allons utiliser la normalisation **Min-Max normalization** pour réduire la plage de variation des amplitudes des signaux d'entrée de manière à ce qu'ils se situent dans l'intervalle $[0,1]$.

Pour tout x_j dans la colonne x :

$$0 \leq \frac{x_j - \min(x)}{\max(x) - \min(x)} \leq 1 \quad (3.1)$$

Et pour tout y_i dans la colonne y :

$$0 \leq \frac{y_i - \min(y)}{\max(y) - \min(y)} \leq 1 \quad (3.2)$$

La Figure 3.7 présente les attributs avant et après la mise en œuvre de la procédure de normalisation.

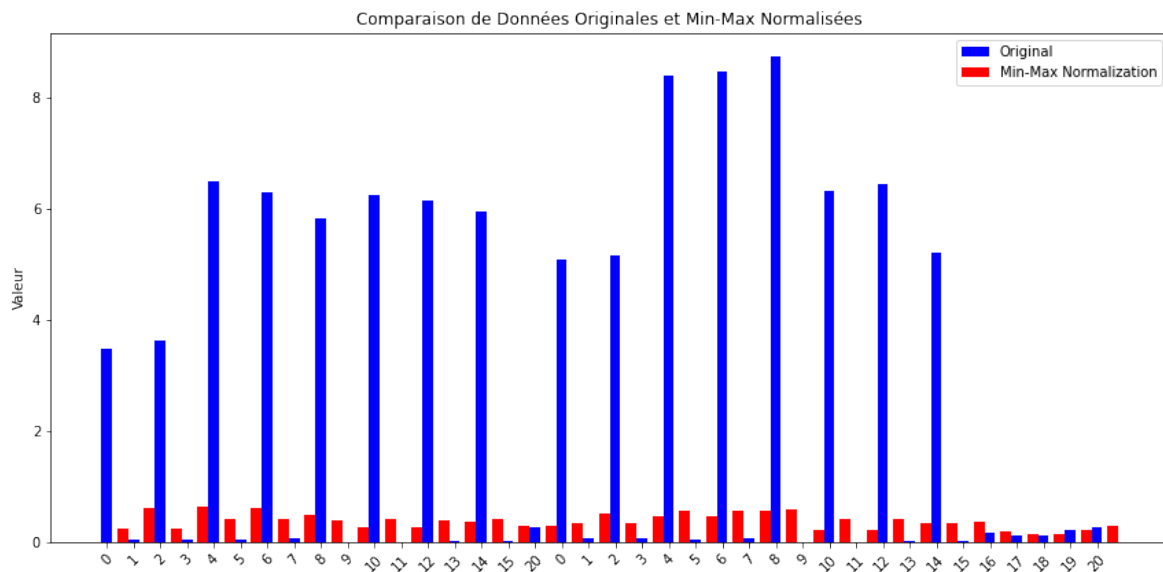


Figure 3.7: Attributs avant et après la normalisation

Une solution pour traiter le problème est de considérer dans le réflectogramme tout le signal temporel, ce qui nécessiterait un nombre de neurones d'entrée important. Pour réduire les entrées, nous n'avons considéré que les pics d'amplitude significative, c'est à dire supérieures à un seuil fixé. Dans le cas, les données d'entrées seront restreintes aux 8 premiers pics significatifs et les instants auxquels ils apparaissent. Par ce processus, nous réduisons considérablement le nombre d'entrée en passant de 2000 échantillons (donc neurones) à 16 pour un réflectogramme donné. Puisque nous effectuerons deux mesures pour une même configuration, nous disposons les mesures vues de R1 et R2 dans la même chaîne d'entrée, ce qui double le nombre de neurones nécessaires.

L'étape suivante de notre processus de prétraitement des données consiste à séparer notre jeu de données en deux sous-ensembles distincts: les données d'entraînement et les données de test.

Les données d'entraînement, souvent notées D_{train} , ce sont celles sur lesquelles nous allons entraîner le modèle. Elles permettent au modèle d'apprendre et de

s'ajuster en fonction des caractéristiques et des tendances présentes dans ces données.

D'un autre côté, les données de test, notées D_{test} , sont utilisées pour évaluer la performance du modèle. Elles offrent la possibilité de vérifier comment le modèle se comporte sur des données qu'il n'a jamais vues auparavant. C'est une étape essentielle pour s'assurer que le modèle est bien généralisable et ne souffre pas de sur-apprentissage. La répartition est de 80% des données pour l'entraînement. Ainsi, pour 1000 échantillons, on aura 800 observations pour l'entraînement et 20%: 200 échantillons pour les tests.

L'étape suivante consiste à créer le réseau de neurones de référence. Pour ce faire, nous utilisons la bibliothèque TensorFlow, pour la construction et la formation du modèle. Dans un premier temps, nous construisons une architecture de type perceptron multicouche (MLP) avec :

- Une couche d'entrée.
- Deux couches cachées, chacune ayant 64 et 32 neurones respectivement, avec une fonction d'activation ReLU.
- Une couche de sortie qui a 5 neurones avec une activation linéaire, ce qui est typique pour les tâches de régression.

Pour l'entraînement nous avons choisi les paramètres suivant:

- Fonction de coût : la fonction de coût utilisée pour entraîner le modèle est l'erreur quadratique moyenne (Mean Squared Error: MSE). Cette fonction est couramment employée pour les tâches de régression, car elle quantifie la différence entre les valeurs prédites par le modèle et les valeurs réelles. Le MSE pénalise davantage les grandes erreurs, ce qui le rend particulièrement utile lorsque nous voulons éviter des erreurs plus importantes dans nos prédictions.
- Optimiseur : l'optimiseur choisi est "Adam". Adam est un algorithme d'optimisation adaptatif des moments d'estimation. Il combine les avantages des deux autres extensions de la descente de gradient stochastique : AdaGrad et RMSProp. Adam est reconnu pour sa capacité à manipuler de grands ensembles de données et des architectures de réseau de neurones avec des paramètres nombreux et/ou des couches profondes.
- Nombre d'époques : le modèle est entraîné pendant 100 époques. Une époque correspond à une itération complète à travers l'ensemble des données d'entraînement. Ce choix garantit que le modèle a eu suffisamment d'itérations pour apprendre et ajuster ses poids permettant de réduire l'erreur

de prédiction.

- Taille de lot (batch size) : Une taille de lot de 32 a été sélectionnée. Cela signifie que 32 échantillons d'entraînement sont utilisés à chaque mise à jour des paramètres du modèle. Cette taille de lot représente un équilibre entre la rapidité d'entraînement (lots plus grands) et la convergence stable (lots plus petits).
- Validation split : 20% des données d'entraînement sont utilisées comme ensemble de validation. Cela permet de suivre la performance du modèle sur un ensemble de données non vu pendant l'entraînement, ce qui est important pour détecter des signes éventuels de surapprentissage.

Outre la fonction de coût MSE, l'erreur absolue moyenne (Mean Absolute Error, MAE) est également utilisée comme métrique pendant l'entraînement. Contrairement au MSE, le MAE donne une idée directe de l'erreur moyenne en unités de la variable cible.

Résultats

Les valeurs des métriques démontrent que le modèle affiche de bonnes performances, en affichant des erreurs MSE (Erreur Quadratique Moyenne), MAE (Erreur Absolue Moyenne) et RMSE (Racine Carrée de l'Erreur Quadratique Moyenne) relativement basses, aussi bien sur les ensembles d'entraînement que de test.

Metric	Training	Test
MSE	0.002	0.003
MAE	0.027	0.030
RMSE	0.050	0.061
R^2	0.970	0.955
MedAE	0.015	0.0169

Table 3.1: Métriques sur l'ensemble de base d'entraînement et de test

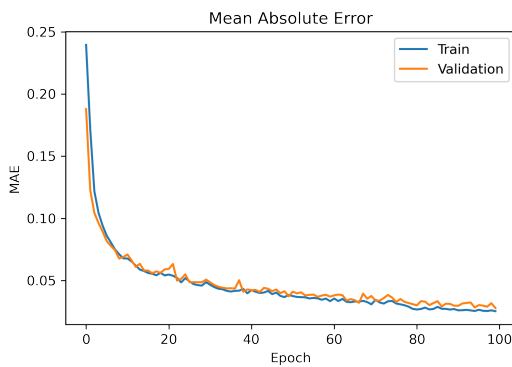


Figure 3.8: Score MAE pour l'ensemble de training vs validation

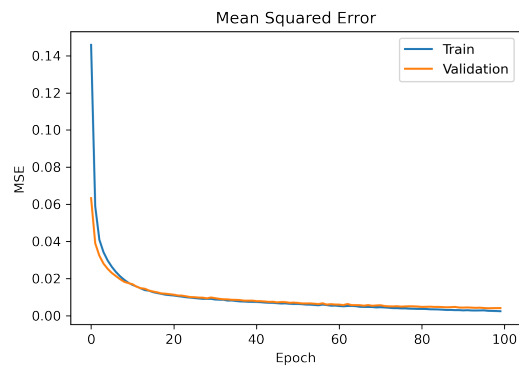


Figure 3.9: Score MSE pour l'ensemble de training vs validation

True Values	Predicted Values	Error
0.875000	0.877327	-0.002327
0.259177	0.276000	-0.016823
0.827004	0.829450	-0.002446
0.394209	0.280822	0.113387
0.371636	0.305641	0.065995
0.857143	0.861148	-0.004006
0.361513	0.408746	-0.047233

Table 3.2: Test Data

3.4.2 Influence des hyperparamètres sur le réseau MLP

Pour effectuer l'optimisation des hyperparamètres, nous nous sommes basés sur la bibliothèque Optuna. Le principal objectif de cette étude est l'analyse de l'influence de chaque hyperparamètre sur notre réseau MLP afin d'obtenir la meilleure combinaison minimisant l'erreur absolue moyenne sur l'ensemble de données de validation. Pour ce faire, nous allons utiliser la fonction objectif qui prend un essai en tant qu'entrée et renvoi la meilleure combinaison de l'ensemble des hyperparamètres suivants:

- Le nombre de couches cachées: on construit notre réseau MLP avec des couches cachées variant de 1 à 10.
- Le choix de l'optimiseur: Trois optimiseurs sont considérés : (Adam, SGD, RMSprop)
- La fonction d'activation: trois fonctions d'activation sont envisagées : ReLU (Rectified Linear Unit), Sigmoid, TanH.
- Le taux d'apprentissage : La recherche pour le taux d'apprentissage est effectuée dans une échelle logarithmique allant de 1×10^{-4} à 0,1.

- Le nombre d'époques: trois durées d'entraînement sont évaluées : 50, 100 et 150 époques.
- La taille du lot (batch size): Des tailles de lot de 16, 32 et 64 sont considérées dans cette étude.

3.4.3 Influence de fonction d'activation

Nous étudions comment la fonction d'activation peut influencer les performances de notre modèle MLP. Nous constatons que les valeurs de MAE avec la fonction d'activation ReLU varient entre 0.0193 et 0.2505, comme le montre la Figure 3.10. Ces scores sont assez bas, indiquant une bonne performance. Concernant les performances avec la fonction d'activation Sigmoid, les scores varient entre 0.0816 et 0.2651. Nous remarquons également que les erreurs augmentent avec le nombre de couches cachées. Les scores MAE varient entre 0.0234 et 0.6341 pour d'autres configurations. En conclusion, la fonction d'activation ReLU offre les meilleures performances globales dans cet ensemble de tests, montrant également une certaine stabilité avec différents nombres de couches.

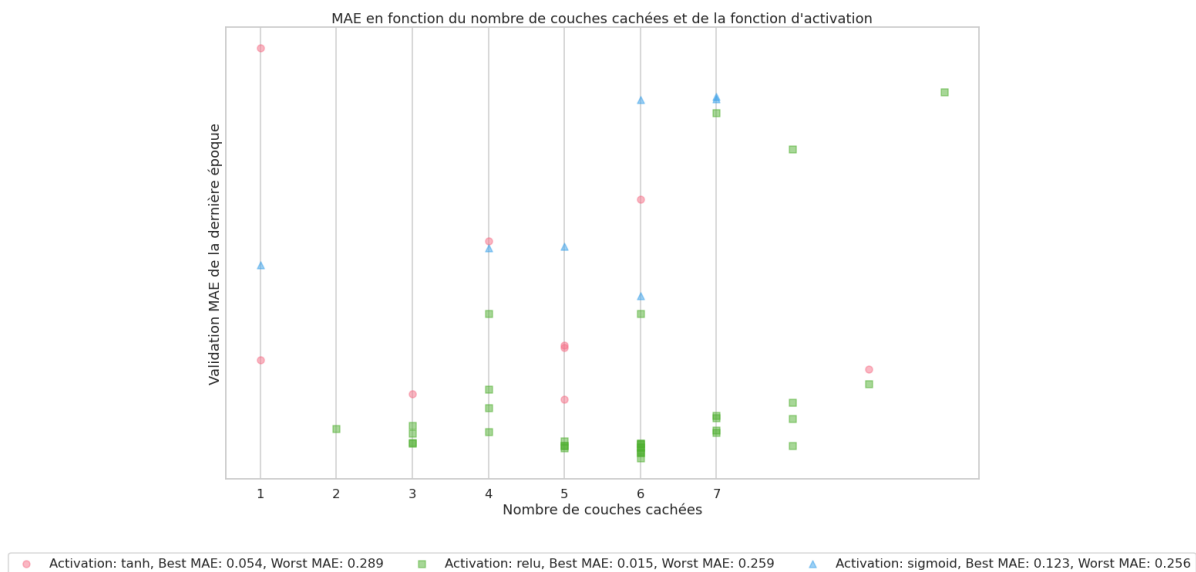


Figure 3.10: Influence de la fonction d'activation

3.4.4 Influence des optimiseurs

Nous observons, comme illustré à la Figure 3.11, que le score MAE varie d'une valeur très basse, 0.0985, jusqu'à 0.2651 pour l'optimiseur SGD. Même en aug-

mentant le nombre de couches cachées (comme 9), le score MAE reste très élevé. Dans le cas de l'optimiseur RMSprop, on remarque que le score MAE varie d'une valeur basse de 0.0234 jusqu'à un pic de 0.6341. De même, lorsqu'on augmente le nombre de couches cachées, le score MAE est très élevé, ce qui peut indiquer un surajustement. Pour l'optimiseur Adam, les valeurs MAE se situent principalement dans la plage inférieure, allant d'une valeur minimale de 0.0193 à un maximum de 0.2505. En conclusion, l'optimiseur Adam semble être le plus performant des trois optimiseurs en termes de MAE minimale.

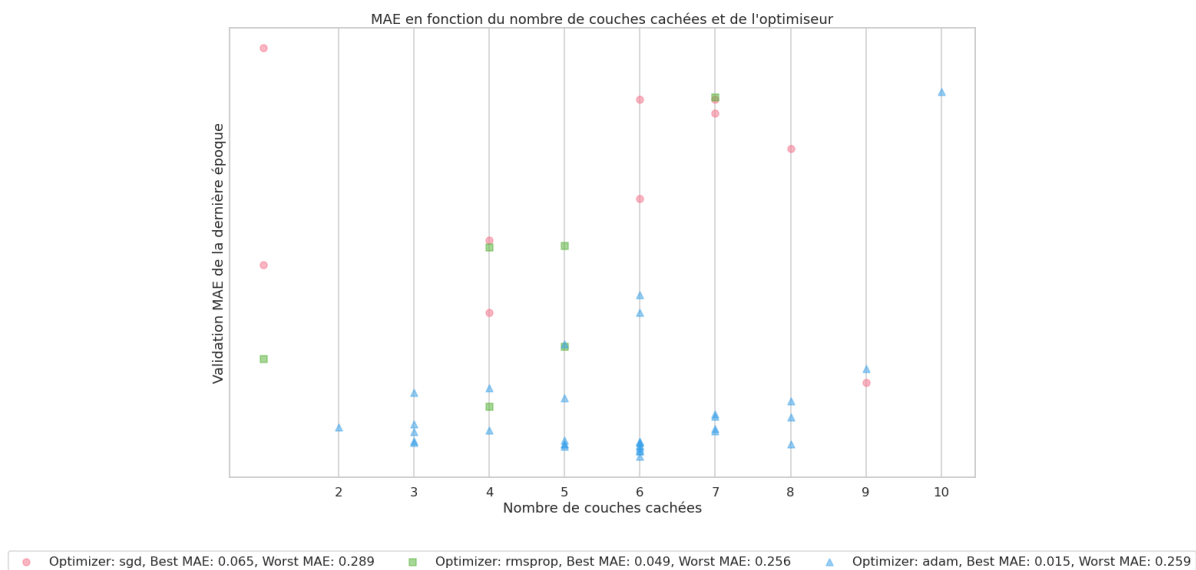


Figure 3.11: Influence des optimiseurs

3.4.5 Influence de nombre de neurones

En analysant les données du tableau, nous pouvons observer que le nombre de neurones par couche et le nombre de couches cachées ont une influence notable sur la valeur finale de la MAE, comme le montre la Figure 3.12. Par exemple, lorsqu'il y a 100 neurones par couche avec 10 couches cachées, la MAE oscille autour de 0.03-0.05. Cependant, lorsque le nombre de neurones est réduit à 70 avec 7 couches cachées, la MAE présente une variabilité plus importante, allant jusqu'à 0.63 dans certains cas.

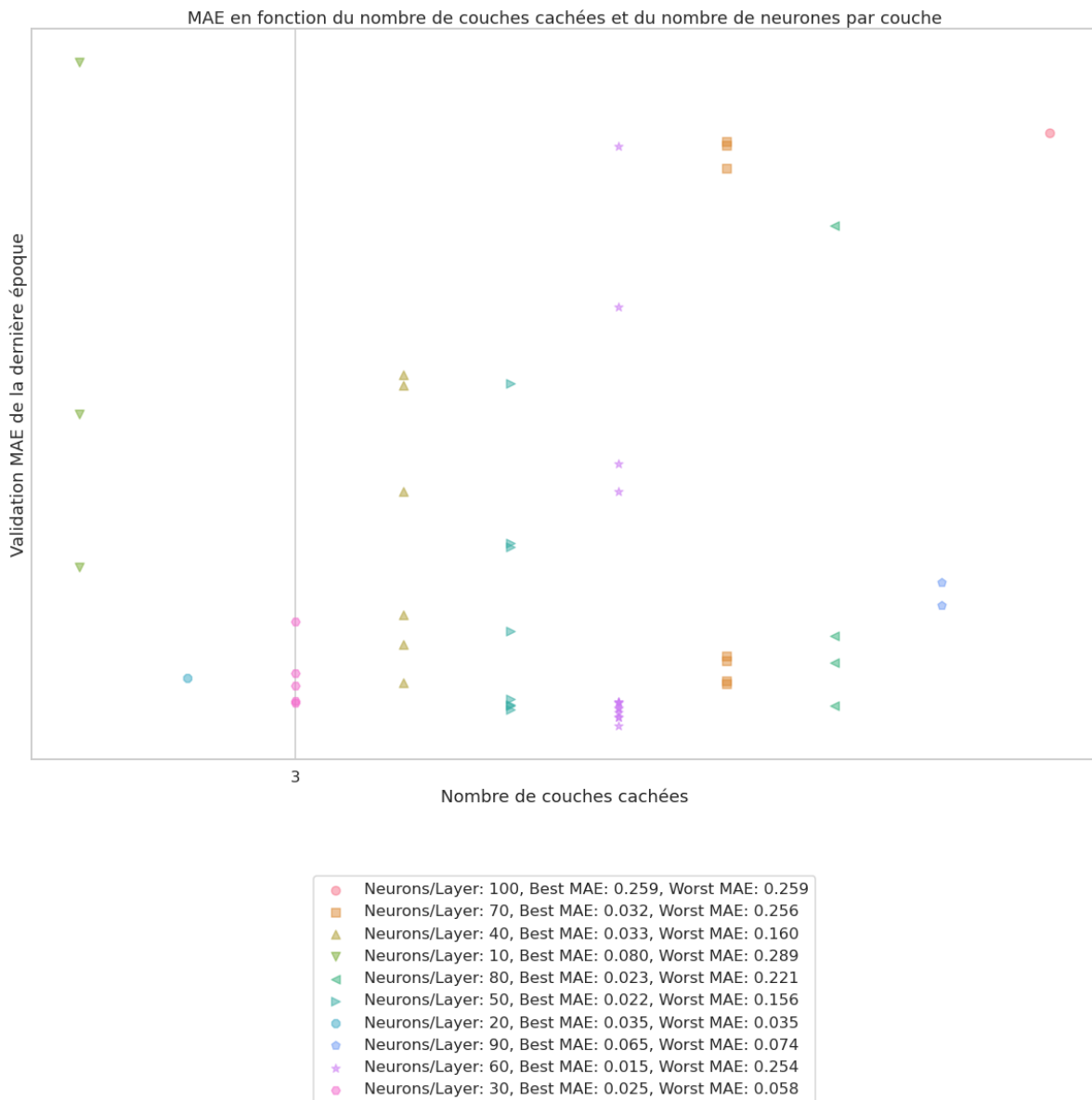


Figure 3.12: Influence du nombre de neurones

Résultats de l'Optimisation Bayésienne:

Hyperparamètre	Valeur
Nombre de couches	5
Fonction d'act.	ReLU
Optimiseur	Adam
Taux d'apprent.	0.001
Nombre d'époques	100
Taille du lot	32

Table 3.3: Meilleurs hyperparamètres pour le réseau MLP.

3.5 Mise en œuvre des réseaux de neurones convolutifs pour la détection des défauts non francs

Dans le cadre de cette étude, nous avons également exploré l'utilisation des réseaux de neurones convolutifs (CNN) pour la détection des positions des défauts. Les CNN, reconnus pour leur efficacité en matière de traitement d'images et de données séquentielles, offrent une approche plus spécialisée pour l'analyse de ce type de données par rapport aux réseaux de neurones classiques. Il est à noter que, pour cette mise en œuvre, nous avons exploité la même base de données que précédemment. De plus, la stratégie de prétraitement appliquée lors de l'expérimentation avec les réseaux de neurones classiques a été conservée sans modification pour les CNN. Ce choix permet d'assurer une comparaison entre les performances des deux types de réseaux et de mettre en évidence les avantages potentiels offerts par les CNN. Les réseaux de neurones convolutifs (CNN) sont traditionnellement utilisés pour traiter les données d'image. Cependant, ils peuvent également être utilisés pour traiter des séquences temporelles. Pour ce faire, nous avons transformé toutes les données séquentielles afin de les rendre compatibles avec notre architecture CNN.

3.6 Sélection des hyper-paramètres optimaux

3.6.1 Influence du nombre des couches de convolution

Nous étudions l'impact de l'augmentation du nombre de couches de convolution sur les performances de notre modèle, comme le montre la Figure 3.13. D'abord, on observe qu'une configuration avec une seule couche de convolution tend à donner de bonnes performances. En effet, de nombreuses occurrences ont des scores d'erreur de validation bas, notamment autour de la valeur 0.20. Cela indique qu'une architecture simplifiée est suffisante pour capturer les caractéristiques essentielles. La performance n'augmente pas nécessairement avec l'ajout de couches supplémentaires; nous observons que, avec deux et trois couches de convolution, le score MAE est notablement moins bon, autour de 0.38.

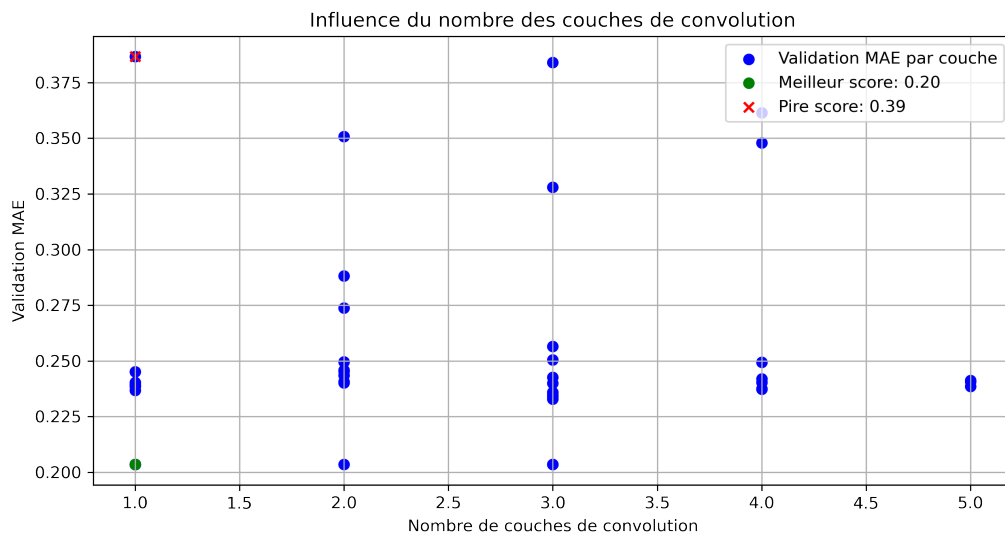


Figure 3.13: Impact du nombre de couches de convolution

3.6.2 Influence du nombre de cachées sur les performances

Nous étudions l'impact de l'augmentation du nombre de couches cachées. Comme on peut le voir dans la Figure 3.14, le meilleur score (MAE le plus bas) semble être atteint avec un nombre de couches cachées de 1 ou de 9, avec des valeurs proches de 0.2035. Le plus mauvais score (MAE le plus élevé) est d'environ 0.3866, obtenu avec 5 couches cachées. On constate que l'augmentation du nombre de couches cachées ne conduit pas à une amélioration directe et constante des performances.

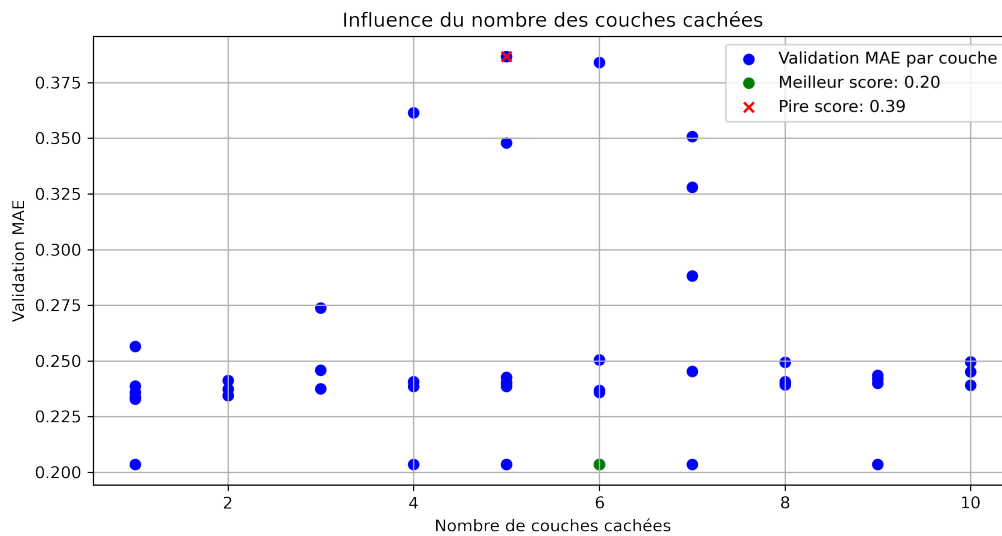


Figure 3.14: Impact du nombre de couches cachées

3.6.3 Influence de la taille des filtres de convolution

Nous étudions l'influence de l'augmentation des tailles des filtres de convolution. D'après la Figure 3.15, nous remarquons que les filtres de taille 8 et 16 ont des performances comparables. Tandis que le filtre de taille 32 offre un bon équilibre entre la performance et la complexité du modèle. Ils ont une performance très stable avec un score MAE autour de 0.23 à 0.25. Pour le filtre de taille 64, les scores sont généralement autour de 0.24, bien qu'il y ait quelques anomalies comme 0.273 et 0.361. Les filtres de taille 128 montrent une grande variabilité dans les performances; nous avons des valeurs comprises entre 0.2035 et 0.3866. Cela pourrait indiquer que les filtres de cette taille peuvent être trop complexes pour le modèle, entraînant à la fois un surajustement (overfitting) et un sous-ajustement (underfitting).

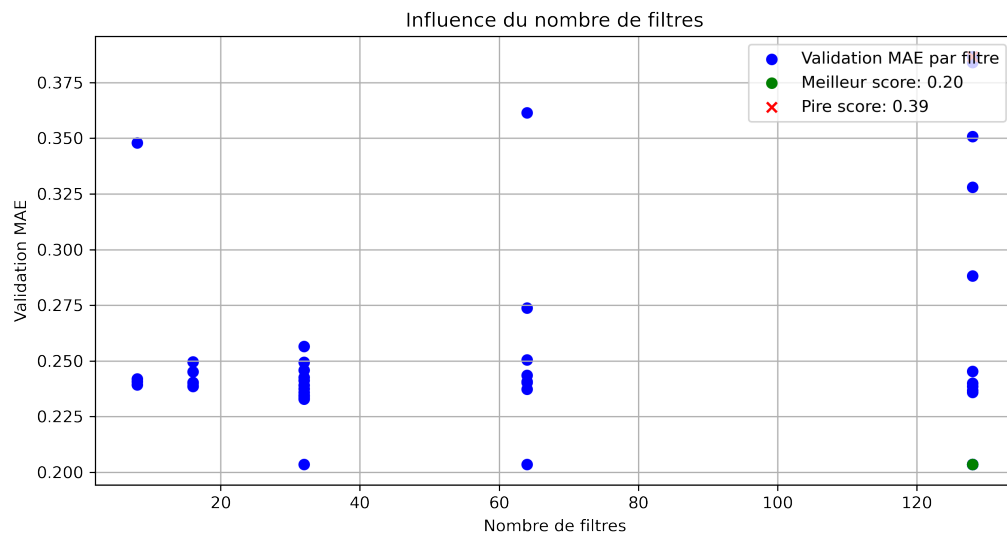


Figure 3.15: Impact de la taille des filtres de convolution

Après avoir effectué une optimisation de nos hyperparamètres, nous sommes parvenus à déterminer les combinaisons les plus performantes. Les résultats suggèrent que les configurations ci-dessous offrent les meilleures performances en termes de précision pour notre modèle:

Le tableau 3.4 présente une valeur de 0.20 avec des paramètres tels que 32 filtres, une taille de noyau de 7, une activation relu, Tandis que le tableau 2 représente une combinaison des hyper-paramètres légèrement différent, notamment pour le nombre de neurones dans chaque couche cachée

Paramètre	Valeurs
Value	0.2035
filters	32
kernel_size	7
activation	relu
optimizer	adam
learning_rate	0.0455
epochs	29
batch_size	92
n_layers	6
n_conv_layers	1
dropout_rate	0.102
n_units_layer0	88
n_units_layer1	64
n_units_layer2	80
n_units_layer3	35
n_units_layer4	39
n_units_layer5	59

Paramètre	Valeurs
Value	0.2035
filters	32
kernel_size	7
activation	relu
optimizer	adam
learning_rate	0.0484
epochs	30
batch_size	92
n_layers	5
n_conv_layers	1
dropout_rate	0.100
n_units_layer0	86
n_units_layer1	68
n_units_layer2	82
n_units_layer3	37
n_units_layer4	35

Table 3.4: Paramètres optimaux pour Rank 1 et Rank 2 après optimisation

3.7 Synthèse

Ce chapitre a porté sur l'analyse des performances de deux types de réseaux de neurones, un Multi-Layer Perceptron (MLP) et un Convolutional Neural Network (CNN), pour la détection de défauts non francs. Différentes métriques telles que le MSE, MAE, RMSE, R^2 , et MedAE ont été utilisées pour évaluer les performances. Dans la première partie, nous avons optimisé les hyperparamètres du modèle MLP, notamment le nombre de couches, l'optimiseur, la fonction d'activation, le taux d'apprentissage, le nombre d'époques, et la taille du lot. Les résultats ont conduit à la détermination des meilleurs hyperparamètres. Ensuite, nous avons étudié l'influence de la fonction d'activation, révélant que ReLU était la plus performante. Les optimiseurs ont également été comparés, avec Adam comme le plus efficace en termes de MAE.

Nous avons également examiné l'impact du nombre de neurones par couche

cachée du modèle MLP, constatant des combinaisons optimales variées en fonction du nombre de couches.

La seconde partie a exploré l'utilisation des CNN pour la détection des défauts non francs. Les résultats ont montré que des configurations plus simples étaient souvent suffisantes.

En conclusion, ce chapitre a permis de déterminer les meilleurs hyperparamètres pour les modèles et d'analyser l'influence de différents facteurs sur leurs performances, fournissant des pistes pour l'optimisation finale du modèle de détection de défauts non francs.

4

Principes Fondamentaux de l'Électromagnétisme Appliqués aux Radars GPR

Sommaire

4.1	Introduction	115
4.2	Notions d'électromagnétisme appliquée au radar	115
4.3	Comportement de la matière vis à vis du champ magnétique	119
4.4	Comportement de la matière vis à vis du champ électrique	120
4.5	Propriétés caractéristiques des sols	122
4.6	Principe de fonctionnement et caractéristiques d'un radar GPR	123
4.7	Performance du radar GPR et Imagerie Radar	128
4.8	Les enregistrement radar	130
4.9	Les différents modes d'acquisition	133
4.10	Antennes GPR	135
4.11	Méthodes numériques pour la construction de B-scan	141
4.12	Conclusion	148

4.1 Introduction

Le radar à pénétration de sol [71] [72] (GPR pour Ground Penetrating Radar) est un instrument d'exploration géophysique qui permet la détection d'objets enfouis dans le sol en utilisant une technique de réflectométrie électromagnétique. En effet, l'émission de signaux connus et la réception de signaux provenant du sol et issus de changements liés à la superposition de couches géologiques ou à des objets enfouis (tuyaux, murs, ...) est à l'origine de diffractions revenant sur l'antenne réceptrice et captés par le radar. Dans la suite, après avoir décrit les phénomènes électromagnétiques liés à la propagation des ondes, nous présenterons le principe de fonctionnement du radar GPR ainsi que ses composants (antennes, systèmes d'acquisition,...). Nous verrons par la suite son dimensionnement suivant l'application visée. Nous donnerons enfin le principe de la modélisation notamment par la méthode FDTD permettant par la suite de construire la base de données et de réaliser l'apprentissage d'un réseau de neurones.

4.2 Notions d'électromagnétisme appliquée au radar

Les radars GPR sont des radars de type électromagnétique, c'est-à-dire basés sur l'émission et la réception d'ondes radiofréquences. Aussi, avant de préciser les technologies radar, nous allons rappeler quelques notions sur la théorie électromagnétique sous-jacente traduisant les phénomènes physiques et de traduire le comportement de la propagation et de la diffraction d'ondes. De même, nous introduisons les équations constitutives traduisant le comportement des milieux sur la propagation des ondes.

4.2.1 Les équations de propagation des ondes électromagnétiques dans les milieux

Les équations de Maxwell

Les équations de diffusion et de propagation permettent de décrire le comportement des ondes électromagnétiques dans un milieu donné. Elles sont fondamentales pour évaluer l'évolution de la propagation de ces ondes.

Les équations de Maxwell, écrites sous leur forme différentielle, décrivent les relations entre les champs électrique et magnétique. Complétées par les paramètres

caractéristiques des milieux, elles fournissent une description mathématique précise de l'interaction entre les champs électromagnétiques et les propriétés du milieu dans lequel les ondes se propagent.

Les paramètres relatifs au champ électrique comprennent le champ électrique lui-même (représenté par le vecteur E), le champ de déplacement électrique (représenté par le vecteur D), la permittivité du milieu (représentée par ε), la conductivité du milieu (représentée par σ), etc.

De même, les paramètres relatifs au champ magnétique comprennent le champ magnétique lui-même (représenté par le vecteur H), le champ d'induction magnétique (représenté par le vecteur B), la perméabilité du milieu (représentée par μ).

Les équations de Maxwell expriment les relations entre ces paramètres. Elles englobent les équations de Gauss pour le champ électrique, les équations de Gauss pour le champ magnétique, la loi de Faraday, et la loi d'Ampère avec et sans terme de déplacement.

Ces équations mathématiques permettent de décrire comment les champs électrique et magnétique se propagent et interagissent avec les propriétés du milieu, telles que la permittivité et la conductivité.

$$\nabla \cdot \vec{E} = \frac{\rho}{\varepsilon_0} \quad \text{Maxwell-Gauss} \quad (4.1)$$

$$\nabla \cdot \vec{B} = 0 \quad \text{Maxwell-Thomson} \quad (4.2)$$

$$\nabla \times \vec{E} = -i\omega \vec{B} \quad \text{Maxwell-Faraday} \quad (4.3)$$

$$\nabla \times \vec{B} = \mu_0 (\vec{J}_C + \vec{J}_S + \varepsilon i\omega \vec{E}) \quad \text{Maxwell-Ampère} \quad (4.4)$$

où:

- \vec{E} est le vecteur de l'intensité du champ électrique (V/m),
- \vec{B} est le vecteur de l'induction magnétique (T);
- \vec{J}_S est le vecteur de la densité du courant électrique (A/m^2).
- ∇ représente l'opérateur différentiel Nabla

Nous pouvons distinguer deux catégories d'équations de Maxwell:

- les équations rotationnelles qui couplent le champ électrique \vec{E} et induction magnétique \vec{B} , constituent les équations de Maxwell-Faraday (MF) et Maxwell-Ampère (MA)
- les équations de la divergence qui ne combinent pas le champ électrique \vec{E} et magnétique \vec{B} , ce sont les équations de Maxwell-Gauss (MG) et Maxwell-

Thomson.

Le principe de l'équation de Maxwell-Thomson, aussi connue sous le nom de loi de Gauss en magnétostatique, énonce qu'il n'existe pas de monopoles magnétiques isolés.

L'équation de Maxwell-Faraday, en rapport avec la règle d'induction de Faraday, permet de voir comment la variation d'un champ magnétique peut engendrer un champ électrique.

L'équation de Maxwell-Ampère, qui correspond au principe d'Ampère en régime variable, prévoit une double création de champs magnétiques: par des courants électriques (c'est le théorème d'Ampère) et par la variation d'un champ électrique. En se plaçant en régime harmonique et en combinant les deux équations en rotationnel dans un milieu homogène sans sources, on peut facilement établir les équations de propagation sur les champs électrique et magnétique comme suit:

$$\nabla^2 \vec{E} - \gamma^2 \vec{E} = 0 \quad (4.5)$$

$$\nabla^2 \vec{B} - \gamma^2 \vec{H} = 0 \quad (4.6)$$

γ désigne la constante de propagation, elle est généralement représenté par l'équation suivante:

$$\gamma = \sqrt{i\omega\mu(\sigma + i\omega\varepsilon)} \quad (4.7)$$

Une solution de cette équation est de type onde plane. L'équation illustrée ci-dessous représente l'expression analytique de celle-ci pour une onde progressive directe. Elle est exprimée comme suit:

$$E = E_0 \cdot e^{-\gamma z} \cdot e^{i\omega t} \quad (4.8)$$

avec: γ représente la constante de propagation exprimée comme suit:

$$\gamma = \alpha + i\beta \quad (4.9)$$

- La partie réelle α représente l'atténuation de l'onde électromagnétique dans le milieu, cette atténuation linéique est représentée par un nombre réel positif

et exprimée en $[Np/m]$. Elle s'exprime mathématiquement par:

$$\alpha = \omega \sqrt{\frac{\mu\epsilon'}{2} \left(\sqrt{1 + \left(\frac{\sigma}{\omega}\right)^2} - 1 \right)} \quad (4.10)$$

- La partie imaginaire β traduit la vitesse de phase de l'onde électromagnétique dans le milieu, elle est exprimée en $[rad/m]$ et s'exprime mathématiquement par:

$$\beta = \omega \sqrt{\frac{\mu\epsilon'}{2} \left(\sqrt{1 + \left(\frac{\sigma}{\omega}\right)^2} + 1 \right)} \quad (4.11)$$

La vitesse à laquelle une onde électromagnétique se propage est associée à la fois à la vitesse de phase dans le vide et aux caractéristiques particulières du milieu dans lequel elle se propage.

$$\nu = \frac{\omega}{\beta} \quad (4.12)$$

L'expression de la vitesse de propagation d'une onde peut être donnée :

$$\nu = \frac{c}{\sqrt{\mu'\epsilon' \left(1 + \left(\frac{\sigma}{\omega}\right)^2 \right)}} \quad (4.13)$$

avec : c représente la célérité d'une onde électromagnétique dans le vide :

$$c = \frac{1}{\sqrt{\mu_0\epsilon_0}} \quad (4.14)$$

$$c \approx 3 \times 10^8 \text{ m/s} \quad (4.15)$$

Dans le cas d'un sol on peut souvent négliger les phénomènes magnétiques la perméabilité relative est alors de 1. De plus, les pertes sont généralement faibles, une expression simplifiée de la vitesse de phase :

$$\nu \simeq \frac{1}{\sqrt{\mu_0\epsilon_0\epsilon_r}} \quad (4.16)$$

Paramètres électromagnétiques de la matière

Ces paramètres jouent un rôle crucial dans la caractérisation et la modélisation du comportement macroscopique d'un milieu lorsqu'il est exposé à une onde électromagnétique. Ils déterminent comment le milieu interagit avec le champ et comment il réagit en termes de conductivité électrique, polarisation électrique, magnétisation et réponse globale aux variations du champ.

4.3 Comportement de la matière vis à vis du champ magnétique

Perméabilité diélectrique

La perméabilité magnétique décrit la capacité d'un matériau à agir sur une variation du champ magnétique, autrement dit à modifier les lignes de flux magnétique. Cette valeur est donc fonction du milieu dans lequel il est produit, où le champ magnétique évolue de façon linéaire, en fonction de l'excitation magnétique pour une valeur de perméabilité supposée constante,

Pour une propagation dans le vide, exempt de toute source externe telle que des aimants, des courants électriques, des matériaux magnétiques ou des conducteurs porteurs de courant, l'expression de la relation entre l'induction magnétique \vec{B} et le champ magnétique \vec{H} peut être formulée comme suit :

$$\vec{B} = \mu_0 \vec{H} \quad (4.17)$$

En régime linéaire et pour un milieu isotrope, le champ magnétique et le champ d'excitation magnétique sont reliés entre eux, dans un matériau donnée, la relation s'exprime comme suit:

$$\vec{B} = \mu \vec{H} \quad (4.18)$$

avec: μ est la perméabilité magnétique du matériau.

4.4 Comportement de la matière vis à vis du champ électrique

Conductivité électrique

En présence d'un champ électrique, deux phénomènes sont observés : le premier est lié à la conductivité électrique, caractérisée par le déplacement des charges électriques libres, et le second concerne la permittivité électrique, qui sera détaillée dans la section suivante.

Lorsqu'on applique un champ électrique à un milieu, on observe un courant de charges. En présence de ce champ, les porteurs de charge se mettent en mouvement, créant ainsi un courant électrique. L'énergie transportée par l'onde se dissipe lors de collisions avec les charges fixes. Cette énergie électromagnétique est alors convertie en chaleur, ce phénomène étant connu sous le nom d'effet Joule.

$$\vec{J}_c = \sigma \vec{E} \quad (4.19)$$

- J_c la densité de courant des charges libres [A/m^2]
- σ représente la conductivité électrique exprimée en [S/m]
- \vec{E} représente le champ électrique exprimée en [V/m]

Permittivité électrique

Quand un matériau est soumis à un champ électrique variable, les particules qui composent ce matériau, et qui sont des charges liées, cherchent à modifier leur position d'équilibre. Cette réaction, appelée polarisation, se produit sur une période minimale appelée temps de relaxation. L'expression du vecteur d'induction électrique est la suivante :

$$\vec{D} = \epsilon \vec{E} = \epsilon_0 \vec{E} + \vec{P}_e \quad (4.20)$$

où:

- \vec{D} : L'induction électrique ($C m^{-2}$)
- ϵ : La permittivité électrique du milieu ($F m^{-1}$)
- ϵ_0 : La permittivité électrique du vide ($F m^{-1}$), avec $\epsilon_0 = 8.854 \times 10^{-12} F m^{-1}$
- \vec{P}_e : La polarisation électrique

La polarisation dépend des caractéristiques des matériaux et des fréquences auxquelles le champ électrique oscille.

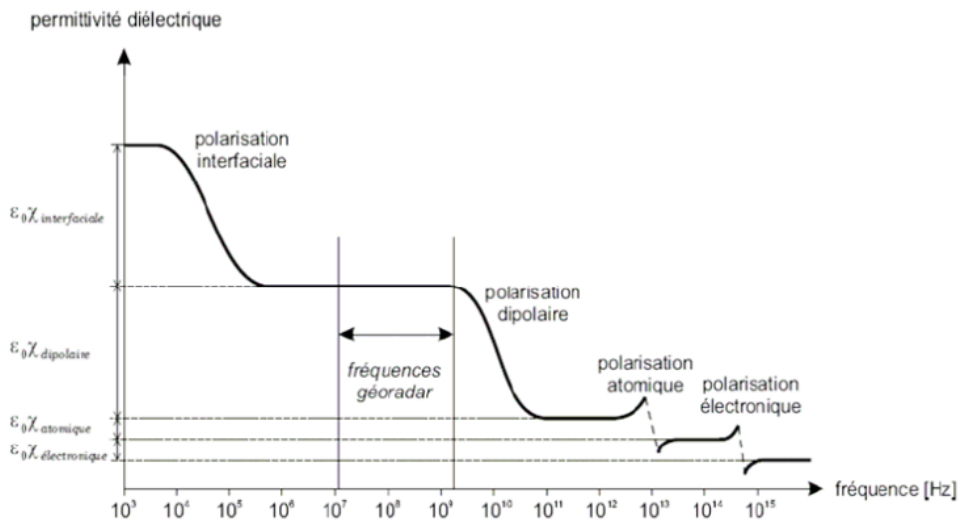


Figure 4.1: Type de polarisation en fonction de la fréquence

Il existe quatre principaux types de polarisation, comme illustré dans la Figure 4.1 :

- La polarisation électronique : elle se produit lorsque le nuage électronique se déforme en réponse à un champ électrique.
- La polarisation dipolaire : dans ce cas, les molécules dipolaires tentent d'aligner leur axe de dipôle avec le champ électrique auquel elles sont soumises.
- La polarisation atomique : cette forme de polarisation se produit dans certains solides où la présence d'ions de charges opposées entraîne des accumulations distinctes en réponse à un champ électrique.
- La polarisation interfaciale : elle se manifeste lorsque des ions s'accumulent au niveau des interfaces dans un matériau hétérogène soumis à un champ électrique.

Dans les bandes de fréquence d'utilisations du radar GPR (comme on peut le voir dans la figure 4.1, la polarisation dipolaire est la plus dominante. Dans les milieux linéaires et isotropes, les relations suivantes s'appliquent :

La polarisation électrique (\vec{P}_e) est liée au champ électrique (\vec{E}) par l'équation suivante :

$$\vec{P}_e = \epsilon_0 \vec{K}_e \vec{E} \quad (4.21)$$

$$\vec{D} = \epsilon_0 \epsilon_r \vec{E} = \epsilon_0 (1 + \vec{K}_e) \vec{E} \quad (4.22)$$

Où :

- ϵ_r représente la permittivité relative.
- \vec{K}_e désigne la susceptibilité électrique.

4.5 Propriétés caractéristiques des sols

La permittivité relative, la conductivité électrique, la vitesse et l'atténuation de divers milieux rencontrés dans les environnements naturels sont présentées dans le tableau 4.1. Cependant, en raison de la présence d'humidité, ces valeurs peuvent varier considérablement dans la bande de fréquence d'un radar à onde de sol. De plus, l'origine de certains matériaux, tels que le granit et le gravier, peut avoir un impact sur les valeurs de permittivité dans une plage de fréquences spécifique, ce qui rend difficile l'obtention de paramètres diélectriques précis. Par conséquent, il est nécessaire d'estimer les caractéristiques diélectriques du sous-sol et de les comparer aux valeurs de la littérature pour chaque site de mesure.

<i>Matériau</i>	ϵ_r	$\sigma(mS/m)$	$\nu(m/ns)$	$\alpha(dB/m)$
<i>Air</i>	1	0	0.3	0
<i>Eau douce</i>	81	0.1-10	0.03	<1
<i>Eau douce glacée</i>	3-4	1	0.15-0.17	<1
<i>Eau de mer</i>	81	4000	0.03	>600
<i>Eau de mer glacée</i>	4-8	10-100	0.11-0.15	8-57
<i>Béton - sec</i>	4-10	1-10	0.09-0.15	<1-5
<i>Béton - humide</i>	10-20	10-100	0.07-0.09	5-36
<i>Granite - sec</i>	5-8	0.001-0.00001	0.11-0.13	<1-5
<i>Granite - humide</i>	5-15	1-10	0.08-0.13	<1-4

Table 4.1: Caractéristiques EM des matériaux mesurées à 100 MHz.

Le tableau 4.1 montre les caractéristiques électromagnétiques de matériaux mesurées à la fréquence de 100 MHz. A cette fréquence, certaines des valeurs typiques incluent la permittivité relative (ϵ_r), la conductivité électrique (σ), la vitesse de propagation (ν) et l'atténuation (α).

4.6 Principe de fonctionnement et caractéristiques d'un radar GPR

Le géoradar est un système d'investigation électromagnétique du sous-sol par émission de signaux électromagnétiques et réception des échos diffractés par les cibles à détecter. Ce type de radar est constitué des éléments suivants:

- Une unité de contrôle,
- Un jeu d'antennes l'une en émission la seconde en réception,
- Un générateur d'impulsions qui fonctionne généralement sur le principe de la décharge de capacités dans le cas de radars fonctionnant en impulsions.
- Un PC servant à l'acquisition et au traitement des données

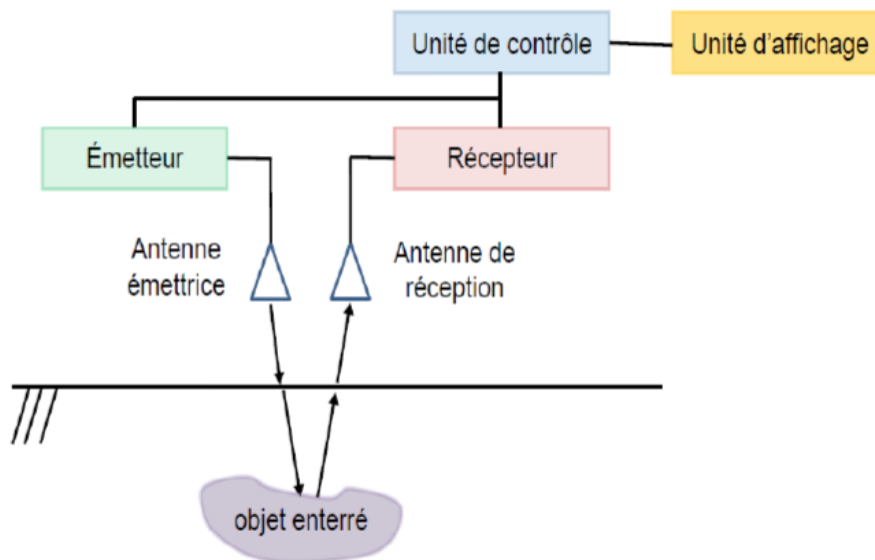


Figure 4.2: Éléments composants un GPR

L'unité de contrôle, servant d'interface entre les différents éléments du GPR, synchronise les étages (voir Figure 4.2). Des fibres optiques ou des câbles coaxiaux relient l'unité de commande à l'électronique d'émission ou de réception tandis que les antennes sont généralement directement connectées à l'électronique. Un convertisseur analogique-numérique et un ordinateur portable permettent une visualisation en temps réel, permettant une interprétation initiale lors de l'acquisition. Au cours du processus, une unité de commande portable commandée par ordinateur régule la transmission via l'antenne d'émission, émettant de façon périodique des impulsions électromagnétiques. Lorsque le radar se déplace à la surface

du milieu, les changements de permittivité diélectrique du sous-sol (comme lors de la présence de roches ou les objets enterrés) donnent lieu à des réflexions. L'antenne de réception capte ces réflexions. Par la suite, le signal reçu subit un traitement numérique post-acquisition pour formater et interpréter les données. L'analyse des données quantitatives est ensuite effectuée pour donner un aperçu de la taille, de la forme, de la profondeur et de l'orientation des objets enfouis dans le sous-sol.

4.6.1 Historique et domaines d'applications du GPR

L'utilisation du radar pour l'analyse du sous-sol a été introduite pour la première fois en 1926 par Hülsenbeck [73]. Malgré ses avantages par rapport aux méthodes sismiques, cette technologie a vu son intérêt diminuer jusqu'aux années 1950. C'est à cette période que l'usage des impulsions électromagnétiques a été relancé, en particulier pour le sondage de la glace profonde. En 1956, El-Said [74] a expérimenté cette méthode pour tenter de mesurer la profondeur de la nappe phréatique à travers des réflexions de signaux. À partir des années 1970, le GPR a commencé à être commercialement employé pour des applications diverses. Les premières utilisations du GPR étaient principalement axées sur le génie civil pour détecter des cavités, des tunnels, des rochers ou encore des câbles et tubes avec des systèmes basés sur des impulsions. Depuis les années 1980, la gamme d'applications du GPR s'est largement étendue. Aujourd'hui, il est couramment utilisé en archéologie, en détection de mines, et pour surveiller l'infiltration de l'eau. D'autres utilisations comprennent l'évaluation de structures en béton, la détection de câbles électriques souterrains, l'étude des interactions sol-racine et la cartographie de l'humidité du sol avec des drones. Chaque application requiert une gamme de fréquence spécifique, adaptée à la taille des objets, la profondeur recherchée et la résolution voulue.

Les systèmes GPR sont essentiels pour détecter de manière non destructive des objets enterrés. Ils sont utilisés dans des secteurs variés tels que le génie civil, l'exploration spatiale, les enquêtes criminelles et lors de catastrophes naturelles [75] [76] [77] [78]. Par exemple, les forces de l'ordre et la police scientifique emploient le GPR pour localiser des objets ou des personnes ensevelies, tandis qu'en génie civil, il permet d'éviter les infrastructures souterraines pendant la construction. De plus, le GPR joue un rôle important dans la détection de mines terrestres [79], [80] [80], contribuant à réduire les dangers liés à ces dispositifs

laissés par les conflits mondiaux. Bien que généralement utilisé en surface, le GPR émet des ondes électromagnétiques depuis le sol pour détecter ces objets enfouis.

4.6.2 Différents types de radar

Les systèmes radar sont classifiés en deux différentes catégories:

Radar Impulsionnel

L'équipement radar moderne utilise une technologie de type impulsion [81]. La particularité de ce système réside dans son processus d'émission et de réception. Une impulsion électromagnétique est émise pendant une durée déterminée, puis le signal de retour est reçu pendant une fenêtre temporelle définie par l'utilisateur. La durée d'impulsion varie généralement entre 0,5 et 10 nanosecondes, selon l'utilisation prévue. Cependant, certaines applications nécessitent des impulsions plus longues pouvant atteindre 1 microseconde, tandis que d'autres facteurs affectent également la durée des impulsions. Pour obtenir des résultats optimaux, l'équilibre approprié doit être trouvé entre la durée d'impulsion, la portée radar et la résolution spatiale, car ces paramètres sont interdépendants.

Décrivons maintenant rapidement les divers composants qui constituent le système radar :

- **Émetteur:** Dans tout système radar, l'antenne joue un rôle central dans la conversion d'un phénomène électrique en un phénomène électromagnétique. Sa fonction principale est de concentrer l'énergie transmise par le système radar dans une direction spécifique préférée par la fonction du système. La composition typique d'un système émetteur implique un commutateur, qui est souvent accompagné d'un système de mise en forme d'impulsions.
- **L'antenne:** La fonction principale d'une antenne est de diffuser de manière optimale les ondes électromagnétiques vers la cible visée avec une perte minimale. Cependant, pour éviter la distorsion du signal, une antenne large bande et peu dispersive est indispensable pour rayonner l'impulsion.
- **Récepteur:** Le système radar comprend un récepteur large bande qui capte le signal réfléchi de la scène. Il s'agit d'un composant critique qui nécessite une sensibilité élevée pour détecter les signaux les plus faibles. Pour des raisons de simplicité, l'antenne réceptrice est souvent identique à l'antenne

émettrice. Généralement, ces deux antennes sont placées côte à côte dans un même support.

Cette technique présente plusieurs avantages, notamment une conception simple et une mise en oeuvre peu coûteuse. Elle garantit également des résultats rapides et faciles à interpréter, car on mesure directement la réponse impulsionnelle sans nécessiter de traitement supplémentaire. Avec son adéquation aux applications embarquées et sa capacité à fournir des résultats complets sur la bande de fréquence en une seule mesure, le système pulsé s'impose comme un choix optimal.

Bien que le radar à impulsions présente de nombreux avantages, il présente également quelques inconvénients à prendre en compte. Le principal défi réside dans le contrôle de la forme de l'impulsion lors de l'émission, car un choix inapproprié peut entraîner une réduction de la résolution. De plus, il est sensible au bruit électromagnétique, ce qui peut entraîner une dégradation de la qualité des données radar.

Un autre inconvénient du radar à impulsions est sa limitation en termes de profondeur de pénétration.

4.6.3 Radar à modulation linéaire de fréquence (Frequency Modulated Continuous Wave)

Les radars FMCW (fréquence modulée à onde continue) sont conçus pour générer un signal à une fréquence variable dans le temps, permettant des ajustements dynamiques de la résolution en fonction de la profondeur et de la nature des cibles. Cela produit un signal sinusoïdal modulé en fréquence appelle "chirp".

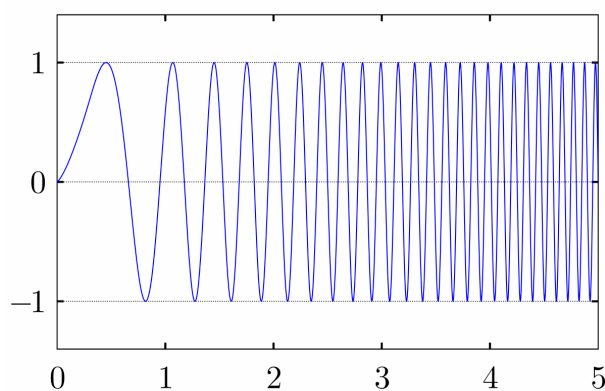


Figure 4.3: Forme d'onde Chirp

La figure 4.3 illustre le principe de fonctionnement d'un radar FMCW. Un VCO, ou oscillateur de contrôle de tension, est utilisé pour générer des fréquences radio en fonction de la tension de commande du DAC (Digital-to-Analog Converter). L'entrée numérique de l'ordinateur de contrôle est transformée en une tension analogique. Le filtre de bande basse reçoit le signal de sortie du VCO, permettant aux fréquences souhaitées de passer. Un coupleur divise le signal en deux: un pour référence dans le mélangeur, et l'autre amplifié pour alimenter l'antenne. L'amplificateur faible bruit traite les signaux reçus transmis via le coupleur. Ces signaux subissent une multiplication, suivie d'un routage vers un filtre passe-bas. Un ADC convertit le signal amplifié résultant en un signal numérique, qui est ensuite envoyé à une unité de traitement pour une analyse plus approfondie.

4.6.4 Autres types de radar

Le GPR est soumis à un large éventail de techniques de modulation et de traitement. Celles-ci incluent :

- La modulation de phase via la modulation de code ou de bruit, ainsi que les méthodes à fréquence unique comme les radiomètres et les approches holographiques.

4.6.5 Le radar RAMAC

Le radar GPR RAMAC commercialisé par la société ABEM est particulièrement apte à détecter des objets enfouis.

En 2003, le département Ondes et Systèmes Associés (OSA) de l'institut Xlim a fait l'acquisition d'un radar Ramac GPR composé de plusieurs composants:

- Un chariot de transport
- Une unité de contrôle CU2
- Trois antennes blindées (250 MHz, 500 MHz, 800 MHz)
- Deux logiciel de traitement (Reflex et Groundvision)



Figure 4.4: Radar GPR RAMAC

La Figure 4.4 illustre un radar RAMAC (photo A), composé d'une antenne blindée de 500 MHz (photo B) et synchronisé par l'unité de contrôle CU2 (photo C). Toutes les antennes sont alimentées par une batterie fournissant un courant continu (photo D).

4.7 Performance du radar GPR et Imagerie Radar

La profondeur de pénétration et la résolution des systèmes de sondage radar [82] [83] caractérisent leurs performances. Celles-ci sont influencées par le milieu sur lequel ils opèrent. La sélection d'un radar dépend de l'application prévue par l'utilisateur, la profondeur de pénétration maximale ou la résolution étant prioritaires en fonction de la situation. Pour aider à sélectionner le radar le plus adapté, diverses études comparatives ont été menées pour estimer les performances de différents systèmes pour des applications spécifiques.

4.7.1 Équation du radar

L'équation du radar permet d'évaluer le bilan des puissances sur le trajet aller-retour d'une onde électromagnétique émise par le système radar. Cette équation

intègre les différents phénomènes physiques qui atténuent la puissance rayonnée en suivant la propagation de l'onde de l'émission jusqu'à la réception du signal réfléchi. Le résultat dépend de plusieurs facteurs, dont les caractéristiques du radar (antennes, circuits électroniques, perte de signal, etc.), et celle de la cible et du milieu de propagation. La puissance nécessaire pour produire un signal et par la suite une image avec un rapport signal sur bruit supérieur à un seuil critique sont déterminées par plusieurs paramètres du système radar tels que la puissance du bruit et le gain de l'antenne. En utilisant l'équation du radar, dans un milieu homogène il est possible d'évaluer la capacité ainsi que les performances du système.

En supposant que la source génère une puissance de P_t , on peut établir que la puissance réellement rayonnée en direction de la cible enfouie vaut:

$$P_{r1} = e_{ot} \cdot G_t(\vartheta_t, \varphi_t) \cdot P_t \quad (4.23)$$

- e_{ot} Ce paramètre représente l'efficacité totale de l'antenne émettrice.
- G_t représente le gain de l'antenne émetteur dans la direction de l'objet enfoui.

La densité de puissance [W/m^2] qui parvient à la cible est égale à:

$$W_1 = P_{r1} \frac{e^{-2\alpha L_c}}{4\pi L_{tc}^2} \quad (4.24)$$

- α : désigne l'atténuation du milieu de propagation
- L_{tc} : représente la distance entre l'antenne émetteur et l'objet enterré.
- P_r : représente la puissance du signal réfléchi reçu par l'antenne réceptrice.
- G_r : représente le gain de l'antenne réceptrice.

La densité de puissance reçue, W_2 , qui retourne à l'antenne réceptrice peut être formulée comme suit :

$$W_2 = \frac{P_r \cdot e^{-2\alpha L_{rc}}}{4\pi L_{rc}^2 \cdot G_r} \quad (4.25)$$

Où W_2 est la densité de puissance à l'antenne réceptrice en [W/m^2].

4.7.2 Dynamique du système

la dynamique du système radar $SP[dB]$ représente le rapport entre la puissance maximale de l'antenne émettrice P_{max} et le niveau de bruit en puissance du

récepteur P_{min} , le facteur de performance $SP[dB]$ peut s'exprimer comme suit:

$$SP = 10 \cdot \log_{10} \left(\frac{P_{max}}{P_{min}} \right) = 20 \cdot \log_{10} \left(\frac{V_{em}}{V_{br}} \right) \quad (4.26)$$

Où: V_{em} représente la tension maximale de l'antenne émetteur et V_{br} représente la tension de bruit de l'antenne récepteur.

4.7.3 Profondeur d'investigation

La profondeur d'investigation, aussi appelée profondeur de pénétration, fait référence à la distance maximale D_{max} à laquelle un système radar peut détecter une cible, en fonction de la puissance maximale émise par l'antenne P_{max} et des caractéristiques spécifiques du milieu. Elle est déterminée en veillant à ce que la puissance du signal reçu par l'antenne soit supérieure à une puissance minimale de référence P_{min} , pour que la cible soit clairement visible malgré les interférences.

$$\frac{P_{min}}{P_{max}} = \frac{1}{p \cdot e_{ot} \cdot e_{ot} \cdot G_t(\vartheta_t, \varphi_t) \cdot G_r(\vartheta_r, \varphi_r) \cdot SP} = S \cdot \lambda^2 \cdot \frac{e^{-4\alpha} e^{D_{max}}}{(4\pi)^3 D_{max}^4} \quad (4.27)$$

4.7.4 Résolution en profondeur

La résolution spatiale représente la distance minimale entre deux objets. Ce paramètre dépend de la durée de l'impulsion, s'exprime comme suit:

$$r = \frac{\nu \zeta}{2} = \frac{\nu}{2B} \quad (4.28)$$

Où ν représente la vitesse de propagation dans le milieu, la division par 2 correspond au trajet aller-retour de l'onde, B correspond à la largeur de bande à mi-hauteur du signal reçu.

4.8 Les enregistrement radar

La technologie GPR offre diverses options de présentation des données, quelle que soit la technologie spécifique utilisée. Le but ultime de la présentation des données GPR est de générer une représentation graphique des données traitées qui reflète avec précision l'image du sous-sol. Il est essentiel de déterminer la

localisation spatiale précise des objets d'intérêt pour faciliter l'interprétation des informations. Dans cette section, nous explorerons trois types de représentation des données utilisées dans GPR.

4.8.1 A-Scan

Un A-scan (appelé aussi une trace) est une représentation unidimensionnelle (1D) de l'amplitude du signal collecté en fonction du temps après avoir placé l'antenne réceptrice au-dessus de la position d'intérêt.

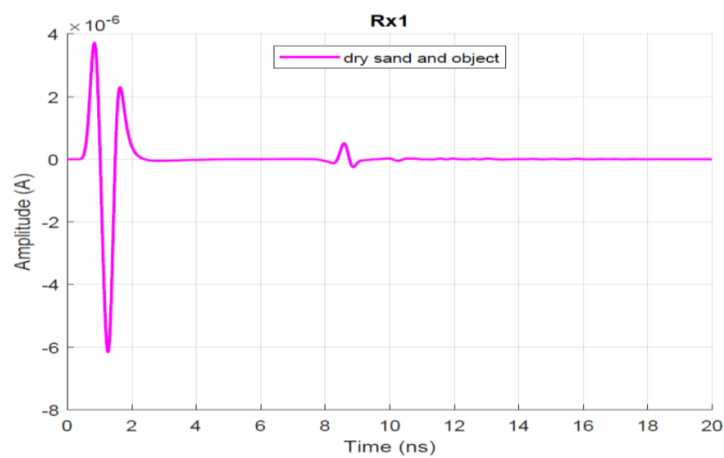


Figure 4.5: Exemple de la forme d'onde d'un A-scan capté par une antenne réceptrice

Un exemple de A-scan est illustré sur la Figure 4.5. Il s'agit du signal capté par l'antenne réceptrice Rx. La première réflexion est souvent un mélange du couplage direct entre les antennes (Tx, Rx1) et la réflexion de la surface du sol. La seconde réflexion, survenant à $t = 8.5$ ns et d'une amplitude moindre, est due à la diffraction de l'objet d'intérêt. Pour améliorer la qualité des enregistrements, il est important d'éliminer à la fois l'écho de surface et le couplage direct entre les antennes, en mettant l'accent uniquement sur le signal d'intérêt. Ces deux interférences peuvent souvent arriver simultanément. Il est possible de gérer cela de deux manières lors de la simulation :

1. Effectuer une mesure en l'absence d'objets.
2. Soustraire cet enregistrement de celui réalisé avec des objets présents dans la scène. Cela permet d'annuler les interférences indésirables.

4.8.2 B-Scan

La méthode A-scan ne permet pas de localiser précisément un objet enfoui. Elle se contente d'indiquer la durée de parcours de l'onde entre les différents objets et l'antenne de réception. Le radar utilisé pour cette méthode est composé de deux antennes proches l'une de l'autre qui se déplacent ensemble. Pour obtenir un B-scan (ou radagramme), plusieurs mesures consécutives sont nécessaires. Le radar doit être déplacé selon une direction spécifique tout en enregistrant des données à des intervalles réguliers le long de cette trajectoire, notée « x ». À chaque position, une mesure est prise. L'ensemble de ces mesures est ensuite compilé pour former le B-scan. Dans les résultats du B-scan, les objets distincts apparaissent sous la forme d'hyperbole. Ces formes proviennent de la mise côte à côte de divers enregistrements pris à différentes positions sur la surface du sol.

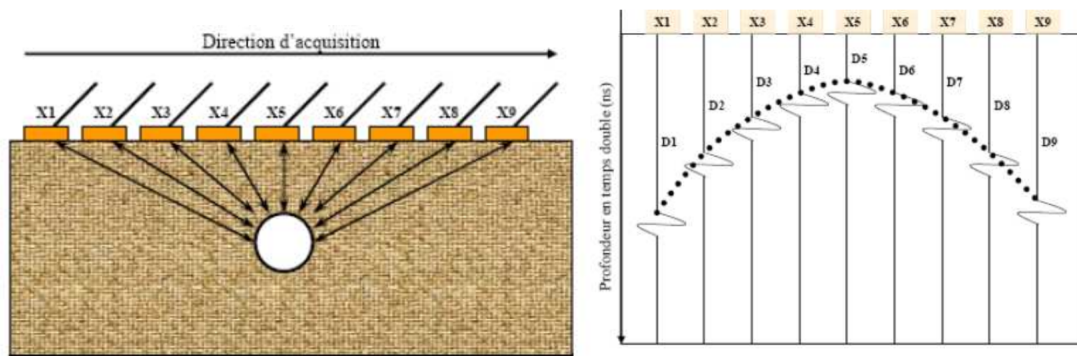


Figure 4.6: Visualisation d'un radagramme: B-Scan.

Le Bscan est souvent utilisé comme mode principal de visualisation pour les enregistrements radar, car il fournit une représentation bidimensionnelle de l'amplitude du signal reçu en fonction de la position et du temps.

4.8.3 C-Scan

Pour acquérir une compréhension plus complète de la géométrie, de la position et de l'orientation d'un objet, plusieurs Bscans sont effectués simultanément; ils sont ensuite fusionnés pour former une image 3D de l'élément souhaité. Ce processus, connu sous le nom de Cscan, Une représentation du type Cscan est illustrée dans la Figure 4.7 ci-dessus.

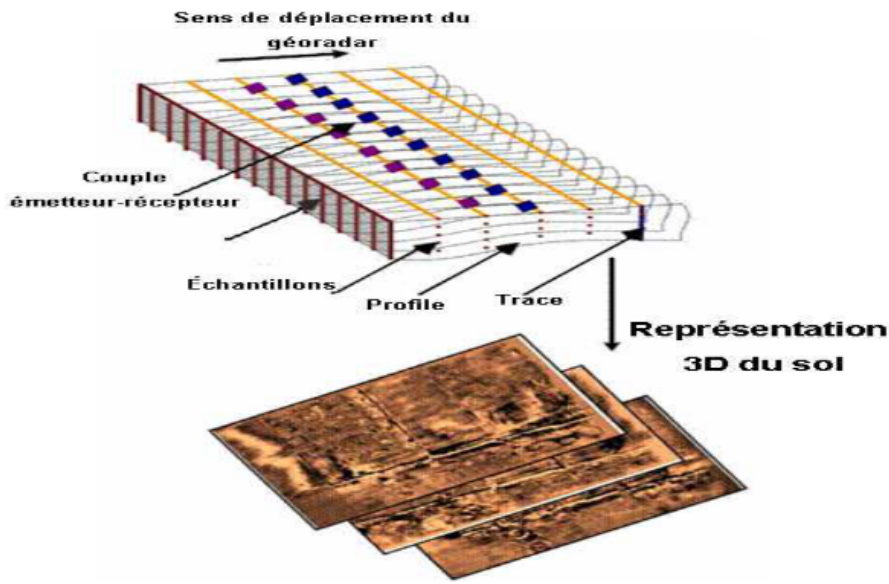


Figure 4.7: Représentation du type Cscan

4.9 Les différents modes d'acquisition

Les mesures GPR peuvent varier en configuration en fonction de l'orientation de l'antenne et du mode de fonctionnement. Chaque configuration unique nécessite des techniques de post-traitement adaptées qui peuvent être utilisées pour optimiser la détection de cibles dans des situations spécifiques. De plus, l'espacement entre les balayages A-Scan dépend de la fréquence centrale (f_c) du signal transmis et des propriétés diélectriques du sol. Pour éviter le crénelage spatial des réponses au sol, le respect de l'intervalle d'échantillonnage de Nyquist est crucial. La taille du pas d'échantillonnage de Nyquist est égale à un quart de la longueur d'onde dans le milieu échantillonné, comme exprimé par l'équation suivante:

$$n_x = \frac{c}{4f_c\sqrt{\epsilon_r}} \quad (4.29)$$

4.9.1 Mode réflexion

Mode à offset constant $Tx - Rx$

La méthode la plus couramment utilisée pour sonder la terre via le radar à pénétration de sol (GPR) est connue sous le nom de mode de décalage commun, elle utilise un seul émetteur et un seul récepteur avec un décalage ou un espace-

ment fixe entre les unités à chaque emplacement de mesure. Les antennes sont déployées selon une géométrie fixe (c.-à-d. séparation, s et orientation) et les mesures sont effectuées à intervalles réguliers, comme le montre la figure 4.7. La majorité des systèmes GPR commerciaux sont construits avec des antennes d'émission et de réception dans une seule enceinte blindée avec une séparation et une orientation fixes. Pour s'assurer que le critère de Nyquist est satisfait, des mesures sont effectuées à intervalles réguliers.

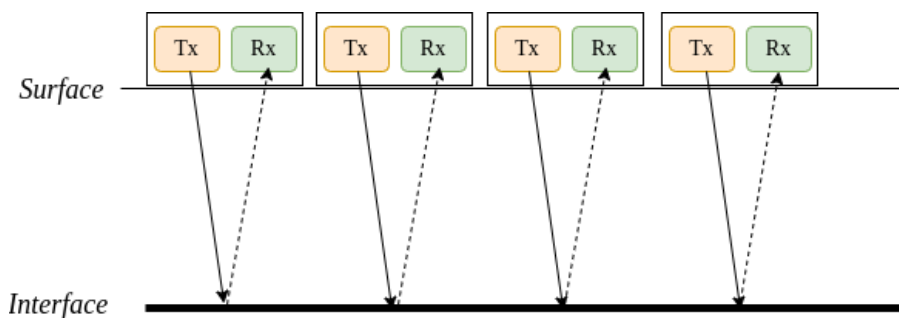


Figure 4.8: Mode avec offset constant : Acquisition en déplaçant les antennes le long d'une ligne horizontale

Mode à offset variable

Dans le mode à offset variable, une technique couramment utilisée dans la géophysique, la distance entre les antennes du radar est progressivement augmentée. Cela signifie que l'écart entre l'émetteur et le récepteur, appelé "offset", change progressivement au cours des mesures. Il existe deux approches courantes pour mettre en œuvre cette variation d'offset :

- Mode WARR: dans cette configuration, l'émetteur reste à une position fixe, tandis que le récepteur est déplacé à des distances variables par rapport à l'antenne fixe. La distance entre les antennes augmente donc progressivement.
- Common-Mid-Point (CMP): Dans cette configuration, les deux antennes sont éloignées symétriquement par rapport à un point central. Cela signifie que l'émetteur et le récepteur sont déplacés simultanément à des distances égales de part et d'autre du point central.

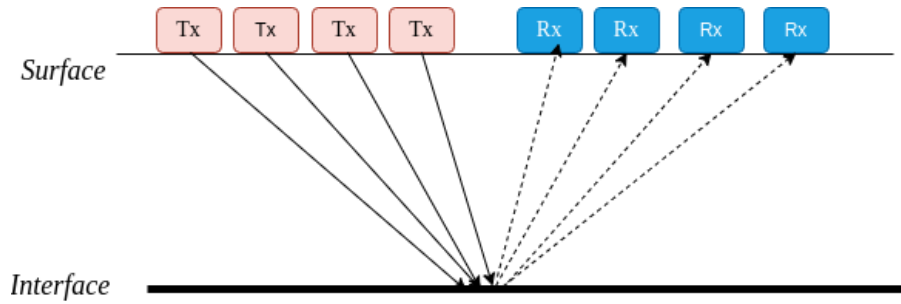


Figure 4.9: Mode avec offset variable

4.10 Antennes GPR

Lors des mesures GPR, l'antenne joue un rôle crucial, établissant la connexion entre le système et l'environnement examiné. Dans cette section, un résumé concis des principaux paramètres influençant le comportement et l'efficacité de l'antenne est présenté. Dans la dernière partie de cette section, les caractéristiques principales et les différentes catégories d'antennes utilisées dans les applications GPR sont explorées.

4.10.1 Directivité et Gain

La directivité d'une antenne représente sa capacité à concentrer le signal dans une direction spécifique. Elle est définie comme le rapport entre la densité de puissance surfacique rayonnée $U(r, \theta_0, \phi_0)$ par l'antenne à une distance donnée et la densité de puissance qui serait rayonnée par une antenne isotrope émettant la même puissance totale P_{ant} .

La directivité est présentée comme suit:

$$D(\vartheta, \varphi) = \frac{U(r, \vartheta, \varphi)}{U_{iso}(r)} = \frac{U(r, \vartheta, \varphi)}{\frac{P_{ant}}{4\pi r^2}} \quad (4.30)$$

La directivité est généralement exprimée en décibels

$$D(\vartheta, \varphi)_{dB} = 20 \log D(\vartheta, \varphi) \quad (4.31)$$

Gain

Le gain d'une antenne est une mesure de sa capacité à concentrer ou à diriger l'énergie rayonnée dans une direction spécifique par rapport à une antenne de référence idéale, appelée antenne isotrope.

$$G(\theta, \phi) \text{ (en dB)} = 10 \log_{10} \left(\frac{P_{\text{rayonnée dans la direction voulue}}}{P_{\text{rayonnée par une antenne isotrope}}} \right) \quad (4.32)$$

Le rendement d'une antenne peut être défini comme la mesure de son efficacité à convertir la puissance électrique qu'elle reçoit ou émet en énergie rayonnée dans la direction souhaitée.

$$\eta = \frac{G(\theta, \phi)}{D(\theta, \phi)} \quad (4.33)$$

4.10.2 Impédance d'entrée de l'antenne

Chaque antenne est caractérisée par son impédance d'entrée (généralement notée Z_e), qui représente un paramètre décrivant la réaction d'une antenne électrique lorsqu'elle est soumise à un signal électromagnétique. L'impédance d'entrée est une quantité complexe comprenant une composante résistive (représentée par R_e) et une composante réactive (représentée par X_e). Soient $V_e(t) = V_e e^{j\phi}$ la tension aux bornes d'entrée de l'antenne et $I_e(t) = I_e e^{j\phi}$ le courant d'alimentation de l'antenne. L'impédance d'entrée de l'antenne est alors définie comme :

$$Z_e = \frac{V_e(t)}{I_e(t)} = R_e + jX_e \quad (4.34)$$

avec :

- R_e : la résistance au point d'entrée de l'antenne.
- X_e : la réactance au point d'entrée de l'antenne.

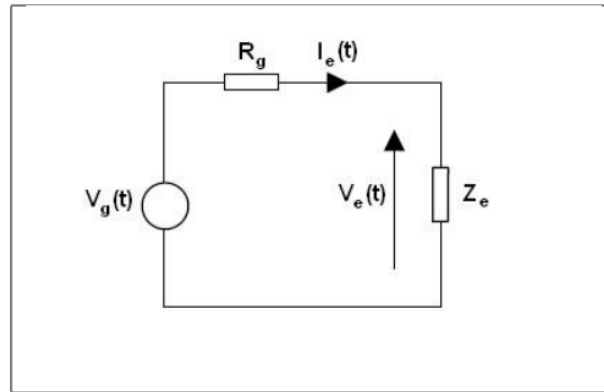


Figure 4.10: Schéma de l'impédance d'entrée de l'antenne Z_e avec alimentation par un générateur

La résistance R_e peut être décomposée en deux composants distincts : la résistance liée au rayonnement de l'antenne, notée R_{ray} , et la résistance liée aux pertes de conduction et diélectrique de l'antenne, notée R_L . La résistance de rayonnement, R_{ray} , représente une résistance fictive qui absorberait toute l'énergie rayonnée par l'antenne sous forme de chaleur.

La puissance alimentant l'antenne se divise donc :

- Une puissance active : $\frac{R_e I_e^2}{2}$
- Une puissance réactive : $\frac{X_e I_e^2}{2}$

4.10.3 Bande passante

La bande passante, généralement notée Δf , désigne la plage de fréquences sur laquelle un signal ou une onde électromagnétique transporte de l'information ou est active. Elle est souvent déterminée par les fréquences de coupure à -20 dB du spectre du signal, notées f_b (fréquence de coupure basse) et f_h (fréquence de coupure haute). La fréquence centrale, souvent notée f_c , est calculée comme étant la moyenne arithmétique des fréquences de coupure, c'est-à-dire :

$$f_c = \frac{f_b + f_h}{2} \quad (4.35)$$

La largeur de bande L_b est définie par la différence entre la fréquence de coupure haute et la fréquence de coupure basse :

$$L_b = f_h - f_b \quad (4.36)$$

4.10.4 Polarisation d'une antenne

La polarisation d'une antenne dans une direction spécifique est déterminée par la manière dont l'onde électromagnétique qu'elle émet est orientée. Selon le théorème de réciprocité, la polarisation en mode de réception est la même que celle en mode d'émission.

On distingue trois types de polarisations différentes:

- **Linéaire** : Linéaire : Le champ électrique oscille dans une direction constante, tandis que le champ magnétique oscille perpendiculairement à ce dernier. Ils sont en phase et oscillent simultanément.
- **Circulaire** : Dans le cas d'une onde transverse électromagnétique (TEM), les composantes du champ électrique ont deux composantes de même amplitude et oscillant en quadrature de phase, c'est-à-dire avec 90 degrés de déphasage entre elles. Les composantes du champ magnétique présentent également ces caractéristiques. Lorsqu'on visualise ensemble ces composantes dans le temps, elles forment une spirale circulaire, illustrative d'une onde polarisée circulairement.
- **Elliptique** : Les caractéristiques sont identiques à une polarisation circulaire mais avec des longueurs de vecteur champ électrique suivant deux axes orthogonaux sont différentes et sont déphasées de 90 degrés. Le champ électrique dessine une ellipse dans le plan de polarisation.

La Figure 4.11 montre des illustrations des trois différents types de polarisation : linéaire, circulaire et elliptique.

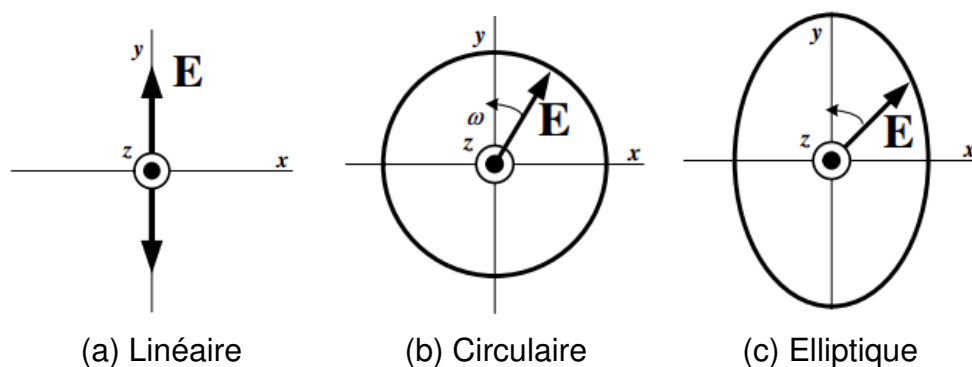


Figure 4.11: Illustration des trois types de polarisation : linéaire, circulaire, et elliptique.

4.10.5 Principaux types d'antennes GPR

Dans cette section, nous introduisons les principaux types d'antennes qui sont utilisées pour les applications typiques de radar géologique. Ce type de radar dispose d'une bande passante assez large. Nous présentons les antennes large bande directive (antennes cornets, antennes papillons), ensuite nous présentons les antennes à transitions progressive (vivaldi), et à la fin de section, nous détaillant ensuite le dipôle Wu-king.

Les antennes sont parmi les éléments les plus critiques d'un radar de pénétration du sol, car elles remplissent les fonctions essentielles de transfert de l'énergie électromagnétique vers la scène étudiée et varie suivant la configuration, la largeur de bande et l'efficacité requises, et de réception de l'énergie diffusée et réfléchi par l'environnement.

Antennes cornets

Les antennes directionnelles à large bande [84] sont classées dans la première catégorie des antennes cornets. Les antennes sont conçues pour émettre et recevoir des ondes électromagnétiques qui pénètrent dans le sol et se réfléchissent sur les objets enterrés.

Une antenne cornet est un type de guide d'ondes qui est élargi à une extrémité pour produire une transition uniforme d'impédance entre le guide d'onde et son milieu environnant. La section de guide d'ondes peut être construite sous diverses formes, telles que rectangulaire ou carrée, tandis que la partie évasée de l'antenne peut également être conçue avec une diversité de formes telles qu'exponentielle ou linéaire. La figure 4.12 montre deux exemples d'antennes cornet, chacune avec des parties évasées et des sections de guide d'ondes uniques.

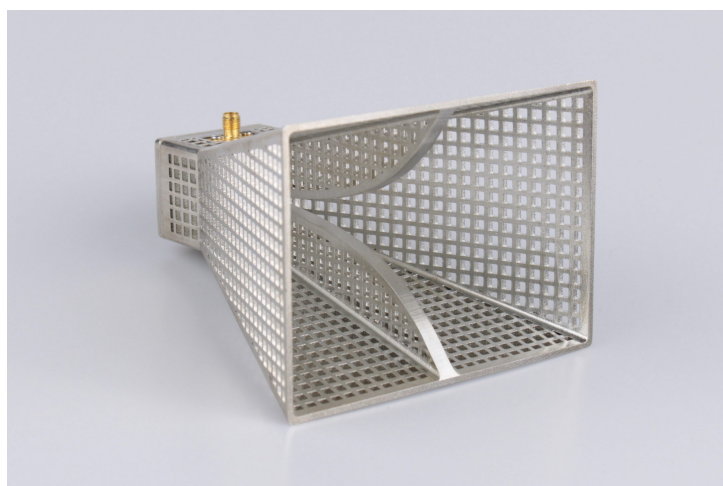


Figure 4.12: Cornet rigide large bande

Antennes papillons

L'antenne papillon [84] s'est révélée être un précieux atout pour divers systèmes radar, notamment les radars GPR. La conception originale de cette antenne comprend deux triangles symétriques par rapport au point d'alimentation (Figure 4.13). Utilisant un guide d'ondes coplanaire pour son alimentation, elle parvient à obtenir une adaptation d'impédance à 50Ω , et à émettre un rayonnement dans toutes les directions, tout en étant économique et compacte.

La fabrication de ces antennes spécifiques utilise la technologie des circuits imprimés. L'incorporation d'une couche résistive aux extrémités des triangles permet à l'antenne de fonctionner en large bande. Afin d'optimiser son rendement tout en conservant la bande passante, un équilibre a été trouvé en associant une charge capacitive à une charge résistive, réduisant efficacement les effets de "ringing" mais au détriment du rendement.

Les antennes de cette conception présentent des réflexions à leurs extrémités; pour les minimiser et élargir la bande passante, des charges résistives sont placées de chaque côté de l'antenne.



Figure 4.13: Antennes papillons.

Antennes Wu-king

Parmi les antennes utilisées dans les systèmes radar GPR, on trouve les antennes WU-KING [85], Ce type d'antenne est défini par une succession de résistances réparties le long de deux brins d'un dipôle électrique. La répartition des charges favorise l'atténuation progressive des courants réfléchis aux deux extrémités du dipôle. L'approche de Wu-King permet une amélioration très significative de la forme d'impulsion rayonnée, et ainsi d'éliminer les signaux bruités.

En raison de leur bande passante ultra-large, de leur petite taille et de l'absence de polarisation croisée, elles sont appropriées pour les applications GPR.

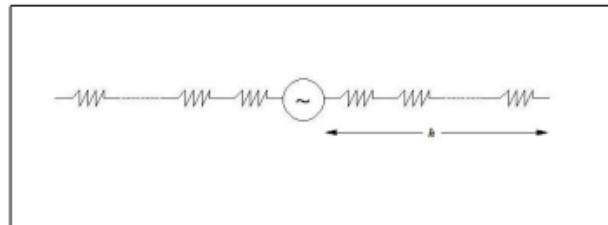


Figure 4.14: Antenne wu king et répartition des charges résistives.

4.11 Méthodes numériques pour la construction de B-scan

4.11.1 Choix de la méthode d'analyse numérique

Les équations de Maxwell peuvent s'écrire sous une forme différentielle comme vu plus haut ou sous une formulation intégrale moyennant le calcul de la fonction de Green relative au milieu de propagation. Pour résoudre numériquement un problème de propagation électromagnétique diverses méthodes associées à ces

formulations ont été développées. L'écriture intégrale permet la construction d'équations intégrales moyennant l'application des conditions limites sur les objets. Ainsi, les inconnues du problème sont alors les courants circulant sur la surface de ceux-ci. Ces courants sont alors décomposés sur les fonctions de base. Moyennant la définition d'un produit scalaire et de fonctions tests, l'application de celui-ci va permettre la construction d'un système matriciel plein. Le problème revient alors à résoudre ce système qui peut être de taille importante si le nombre d'inconnues en courant croît. De plus, nous voyons qu'il est plus difficile de traiter un problème de propagation dans un sol, notamment si celui-ci est inhomogène par une telle formulation. Aussi, nous avons préféré utiliser des techniques de résolution des équations locales aux dérivées partielles de Maxwell. Pour ce faire diverses approches peuvent être utilisées:

La méthode des éléments finis [86] est plutôt une approche utilisée dans le domaine fréquentiel pour l'étude de composants tels des filtres ou des systèmes antennaires. La matrice construite, bien que creuse, peut devenir lourde à résoudre. De plus le traitement fréquence par fréquence n'est pas adapté à l'étude d'une problématique large bande malgré les évolutions pour éviter cet inconvénient. La méthode des volumes finis [87]. Cette approche consiste à réaliser un découpage volumique de l'espace comme la méthode des éléments finis. Toutefois le transfert de l'énergie entre les cellules est diffusif malgré la construction de schémas d'ordre supérieurs. Aussi, nous n'avons pas retenu ce type d'approche plutôt utilisé en mécanique des fluides.

Le méthode TLM (Transmission Line Matrix) [88] est une technique qui consiste à résoudre le problème tridimensionnel en temporel en décomposant le volume en hexaèdres. Le calcul de la propagation des ondes est alors effectué à partir d'une succession d'opérations de propagation entre les nœuds du maillage et de diffraction au niveau de ceux-ci. C'est une approche intéressante pour des applications GPR.

Le méthode des différences finies [89] dans le domaine temporel (FDTD), il s'agit de mailler l'espace en hexaèdres. Le calcul des champs électromagnétiques s'effectue au sein de chaque cellule et le transfert d'une cellule à la suivante est assurée par les équations de Maxwell en rotationnel. La résolution se fait au fil des itérations temporelles et utilise des formes d'onde impulsionnelles. Le traitement de milieux hétérogènes n'entraîne pas de surcoût de calcul. Toutes ces propriétés font de cette méthode une bonne candidate pour notre problématique.

Suite au bilan des méthodes d'analyse d'un problème de propagation en élec-

électromagnétique et compte tenu de la spécificité de la problématique GPR, deux conditions s'imposent:

- La propagation pouvant se faire dans des sols à caractéristiques électromagnétiques variables, une approche 3D semble préférable,
- Les GPR qui feront l'objet de nos calculs fonctionnent en régime impulsionnel, aussi, une approche temporelle est souhaitée

Pour ce faire, nous avons le choix entre la TLM et la FDTD, compte tenu des compétences de l'équipe CEM et diffraction de Xlim qui a développé un code FDTD 3D: TEMSI-FD, c'est cette dernière qui a été retenue dans le cadre de cette thèse.

4.11.2 Principe de la méthode FDTD

La méthode numérique FDTD (Finite Difference Time Domain) [89] a été mise en place par Kane S. Yee en 1966. C'est seulement à partir des années 1975, qu'un nombre de publications sans cesse croissant a révélé ses nouvelles applications et possibilités. Toutefois, la proposition originale de Yee d'une grille cartésienne rectiligne continue de former la base de la méthode FDTD. La méthode implique une discrétisation des équations de Maxwell à la fois dans le temps et dans l'espace. La méthode FDTD partitionne l'espace en cellules de Yee, des parallélépipèdes rectangulaires de dimensions soit identiques, soit non uniformes (avec un pas de maillage variable par direction cartésienne, comme le montre la figure 4.15. Ces cellules ont des propriétés EM homogènes et les composantes des champs électriques et magnétiques sont déterminées en un nombre fini de points. Il est important de noter que, plus la discrétisation spatio-temporelle est fine, plus il y a de cellules créées et donc plus le temps de calcul est long.

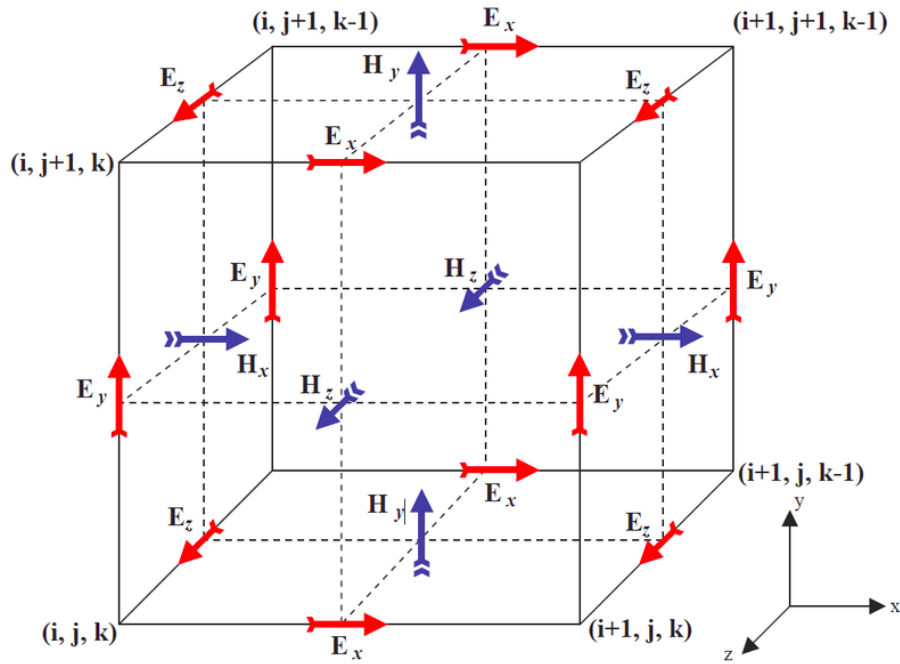


Figure 4.15: Schéma d'une cellule de Yee en 3D.

A travers le développement de Taylor, les différences finies centrées du second ordre fournissent une approximation des dérivées partielles du second ordre par rapport au temps et l'espace sur les composantes de champ . Ceci est illustré dans les équations données ci-dessous:

$$\frac{\partial u}{\partial x}(i, j, k, n) = \frac{u_{i+\frac{1}{2},j,k}^n - u_{i-\frac{1}{2},j,k}^n}{\Delta x} + O[(\Delta x)^2] \quad (4.37)$$

- $\Delta x, \Delta y$ représente les pas spatiaux dans la direction (ox), et (oy) respectivement,
- Δt est le pas temporel et n représente un entier qui désigne l'instant de cellule $t = n\Delta t$.

Dans un milieu diélectrique isotrope et linéaire, Les équations différentielles de Maxwell, projetées dans un repère cartésien (o, x, y, z), s'expriment comme suit:

$$\frac{\partial E_x}{\partial t} = \frac{1}{\varepsilon} \left(\frac{\partial H_z}{\partial y} - \frac{\partial H_y}{\partial z} - \sigma E_x \right) \quad (4.38)$$

$$\frac{\partial E_y}{\partial t} = \frac{1}{\varepsilon} \left(\frac{\partial H_x}{\partial z} - \frac{\partial H_z}{\partial x} - \sigma E_y \right) \quad (4.39)$$

$$\frac{\partial E_z}{\partial t} = \frac{1}{\varepsilon} \left(\frac{\partial H_y}{\partial x} - \frac{\partial H_x}{\partial y} - \sigma E_z \right) \quad (4.40)$$

$$\frac{\partial H_x}{\partial t} = \frac{1}{\mu} \left(\frac{\partial E_y}{\partial z} - \frac{\partial E_z}{\partial y} - \sigma^* H_x \right) \quad (4.41)$$

$$\frac{\partial H_y}{\partial t} = \frac{1}{\mu} \left(\frac{\partial E_z}{\partial x} - \frac{\partial E_x}{\partial z} - \sigma^* H_y \right) \quad (4.42)$$

$$\frac{\partial H_z}{\partial t} = \frac{1}{\mu} \left(\frac{\partial E_x}{\partial y} - \frac{\partial E_y}{\partial x} - \sigma^* H_z \right) \quad (4.43)$$

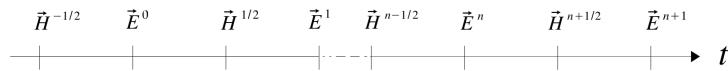


Figure 4.16: Schéma saute-mouton (leapfrog) représentant le principe de mise à jour des champs électrique et magnétique

La Figure 4.16 montre comment chaque composante de champ E dans une cellule Yee est enveloppée par quatre composantes de champ H , et vice versa. Cette configuration spatiale fait que les composantes sont décalées d'une demi-cellule et s'échelonnent dans le temps selon un modèle de saute-mouton. Tout d'abord, les composantes du champ E sont mises à jour à l'aide des composantes du champ H calculées précédemment. Ensuite, les composantes du champ H sont mises à jour à l'aide des composantes du champ E précédents. Enfin, le processus est itératif avec un pas temporel défini.

La discrétisation du champ magnétique en espace et en temps est déterminé comme suit:

$$H_x|_{i,j,k}^{n+\frac{1}{2}} = H_x|_{i,j,k}^{n-\frac{1}{2}} - \frac{dt}{\mu_{i,j,k} dy} \left[E_z|_{i,j+1,k}^n - E_z|_{i,j,k}^n \right] + \frac{dt}{\mu_{i,j,k} dz} \left[E_y|_{i,j,k+1}^n - E_y|_{i,j,k}^n \right] \quad (4.44)$$

$$H_y|_{i,j,k}^{n+\frac{1}{2}} = H_y|_{i,j,k}^{n-\frac{1}{2}} - \frac{dt}{\mu_{i,j,k} dz} \left[E_x|_{i,j,k+1}^n - E_x|_{i,j,k}^n \right] + \frac{dt}{\mu_{i,j,k} dx} \left[E_z|_{i+1,j,k}^n - E_z|_{i,j,k}^n \right] \quad (4.45)$$

$$H_z|_{i,j,k}^{n+\frac{1}{2}} = H_z|_{i,j,k}^{n-\frac{1}{2}} - \frac{dt}{\mu_{i,j,k} dx} \left[E_y|_{i+1,j,k}^n - E_y|_{i,j,k}^n \right] + \frac{dt}{\mu_{i,j,k} dy} \left[E_x|_{i,j+1,k}^n - E_x|_{i,j,k}^n \right] \quad (4.46)$$

La même procédure s'applique également au champ électrique E_x , E_y , E_z .

4.11.3 Stabilité de calcul et dispersion numérique

En analyse numérique, la méthode FDTD est réputée pour être conditionnellement stable. Elle impose une borne supérieure au pas de temps Δt , au-delà de laquelle la solution numérique devient instable. La condition **Courant-Friedrichs-Lewy (CFL)** détermine le critère de stabilité de la méthode, qui s'applique à tous les types de milieux diélectriques. Selon cette condition, le pas de temps Δt doit être restreint par les pas spatiaux Δx , Δy , Δz pour assurer la stabilité de la solution numérique.

$$\Delta t \leq \Delta t_{max} = \frac{1}{\nu \sqrt{\left(\frac{1}{\Delta x}\right)^2 + \left(\frac{1}{\Delta y}\right)^2 + \left(\frac{1}{\Delta z}\right)^2}} \quad (4.47)$$

Le critère de stabilité est basé sur un milieu homogène et sans perte, où la vitesse de propagation des ondes électromagnétiques est notée ν . En réalité, les modèles présentent souvent des régions de propriétés et de vitesses variables. Dans la modélisation GPR, la présence d'espace libre nécessite que ν soit défini comme la vitesse de la lumière. Lorsque $\Delta x = \Delta y = \Delta z = \Delta l$, le critère de stabilité s'exprime par:

$$\Delta t \leq \frac{\Delta l}{c\sqrt{3}} \quad (4.48)$$

La méthode FDTD est sujette aux erreurs, dont l'une est la dispersion numérique, comme mentionné dans de nombreux travaux [90] [91]. L'utilisation des différences finies comme maillage de domaine de calcul pour la propagation des ondes EM peut provoquer des distorsions. Ces effets sont principalement attribués à la discrétisation et au calcul des dérivées partielles de la méthode de calcul. Pour minimiser les problèmes potentiels causés par la dispersion numérique, il est recommandé de choisir un pas de discrétisation spatial qui est au moins dix fois plus petit que la plus petite longueur d'onde des champs électromagnétiques propagés:

$$\Delta l \leq \frac{\lambda}{20} \quad (4.49)$$

4.11.4 Modélisation de l'espace infini

Le rayonnement des antennes ainsi que les ondes diffractées par les objets se propagent dans le sol ou l'air dans un milieu de grande dimension que l'on peut considérer comme semi infini. Or, d'un point de vue numérique, de par la nécessité de discrétiser l'espace environnant les objets étudiés (antennes, tuyaux ..), l'espace représenté est nécessairement de dimension finie. Aussi, il est nécessaire de disposer sur le pourtour du volume de calcul des conditions permettant d'absorber les ondes de manière à ce qu'elles ne reviennent pas polluer le volume utile. Pour ce faire, historiquement, des conditions d'absorption résolvant des équations d'advection ont tout d'abord été utilisés puis l'introduction d'un milieu absorbant à pertes à très vite pris le dessus avec l'introduction en 1994 des PML (Perfectly Matched Layers) par Jean Pierre Berenger. En effet, à la manière d'une chambre anéchoïde, le milieu PML permet une adaptation théorique parfaite en impédance et une atténuation de l'onde se propageant dans celui-ci. Cette opération est réalisée en définissant un milieu possédant à la fois une conductivité électrique et une conductivité magnétique. Ceci permet une adaptation en impédance de l'onde quelle que soit sa direction d'incidence sur l'interface entre l'air et la PML. Le choix judicieux des valeurs de celles-ci et de la loi de variation (en puissance ou géométrique) permet de répondre de façon très satisfaisante au problème puisqu'on peut atteindre des coefficients de réflexion numérique très faibles pouvant descendre en dessous de -100dB. Dans la pratique, on limite le nombre de couches PMLs à une dizaine pour un coefficient de réflexion compris entre -60dB et -80dB.

4.11.5 Modélisation des antennes filaires

Ce modèle repose sur celui développé à l'origine par R. Holland (publi Holland & Simpson). Il consiste à ajouter deux équations couplées sur le courant et la tension (ou la densité de charge) à la manière d'une théorie de lignes de transmission pour ligne bifilaire. Il consiste à ajouter deux équations couplées sur le courant et la tension (ou la charge) à la manière d'une théorie de lignes de transmission pour ligne bifilaire. La principale différence provient du fait que le retour se fait par la maille à la manière d'un coaxial traduisant le fil dans la maille. Ainsi, les self et capacités seront non plus référencées par une masse mais par la maille fictive entourant le fil. Un couplage bilatéral réciproque permet la prise en compte du couplage entre le courant circulant sur le fil et le champ autour de celui-ci. Ce

modèle a été largement validé et amélioré en une version fils obliques notamment, il ne fera pas l'objet de développement particulier dans ce mémoire.

4.12 Conclusion

Ce chapitre établit les bases essentielles pour comprendre la technologie du radar à pénétration de sol (GPR) en tant que méthode d'exploration géophysique. Nous avons exploré les principes fondamentaux du GPR, y compris les phénomènes électromagnétiques liés à la propagation des ondes et les composants clés du système GPR, tels que les antennes et le système d'acquisition.

De plus, nous avons discuté du dimensionnement du GPR en fonction des applications spécifiques et de la modélisation de son fonctionnement à travers la méthode des différences finies dans le domaine temporel (FDTD), une approche adaptée aux applications GPR. Ces bases sont ensuite utilisées pour le développement d'une base de données permettant de générer des images synthétiques.

En outre, nous avons abordé les différents types de systèmes radar, en mettant en évidence les avantages et les inconvénients des systèmes radar impulsionnels, qui émettent des impulsions électromagnétiques pour la détection.

Pour finir, nous avons évoqué les méthodes de résolution numérique des équations de Maxwell, notamment la méthode des différences finies dans le domaine temporel (FDTD), qui est une approche efficace pour la simulation en temps réel des problèmes de propagation électromagnétique en temporel.

Ce chapitre établit les concepts fondamentaux du GPR, qui seront explorés plus en détail dans les sections suivantes de ce travail, notamment pour le développement de la base de données et la génération d'images synthétiques.

5

Contribution et développement des architectures de l'apprentissage profond pour la classification des images B-scan

Sommaire

5.1	Introduction	151
5.2	Création de la base de données	151
5.3	Modélisation FDTD pour la création de la base de données	157
5.4	Tracé des B-scans et recalage temporel	158
5.5	Prétraitement des données	161
5.6	Détection des contours	163
5.7	Architectures utilisées pour la multi label classification des images B-Scans	169
5.8	Architecture VGG16	177
5.9	CNN personnalisé	180

Chapitre 5 – Contribution et développement des architectures de l'apprentissage profond pour la classification des images B-scan

5.10 Influence des hyperparametres	182
5.11 Synthèse	189

5.1 Introduction

Ce chapitre a pour objectif d'expérimenter l'adaptation des réseaux de neurones profonds dans le domaine des radars à sondage de sol. Ainsi, nous déroulerons tout le processus permettant de réaliser une classification d'objets enfouis en partant de l'image synthétique obtenue théoriquement par FDTD. Dans un premier temps, nous détaillerons la base de données générée pour la démonstration, puis le prétraitement des images obtenues sera effectué pour permettre l'apprentissage de réseaux profonds. Ensuite, nous mettrons en oeuvre trois architectures que nous avons développées et comparerons leurs performances. Ces dernières seront analysées en utilisant des métriques adaptées à la problématique de classification. Enfin, nous testerons le pouvoir prédictif des modèles ainsi mis en place sur des échantillons n'appartenant pas à la base de données d'apprentissage.

5.2 Création de la base de données

5.2.1 Configuration Multi-Capteurs

Pour construire notre base de données, nous allons nous focaliser sur des configurations pour lesquelles l'émetteur est fixe et le récepteur se déplace suivant une direction fixe sur l'interface air /sol. Ainsi, l'antenne émettrice sera située sur la gauche de la scène et l'antenne de réception va se déplacer d'un pas $dx = 20$ mm. Alors, si nous voulons réaliser un déplacement sur une distance de 0.78 m, cela nécessite 40 simulations pour générer tous les A-scans nécessaires à la création du B-scan. Aussi, nous avons considéré un réseau de $N = 10$ antennes séparées de ce pas dx . Ainsi en déplaçant le réseau d'un pas égal à $(N + 1) \cdot dx$, cela permet de diviser par N (c'est-à-dire 10 dans le cas présent) le nombre de simulations à effectuer. Ce processus est réalisable en supposant les antennes faiblement couplées entre elles, ce qui a été préalablement vérifié.

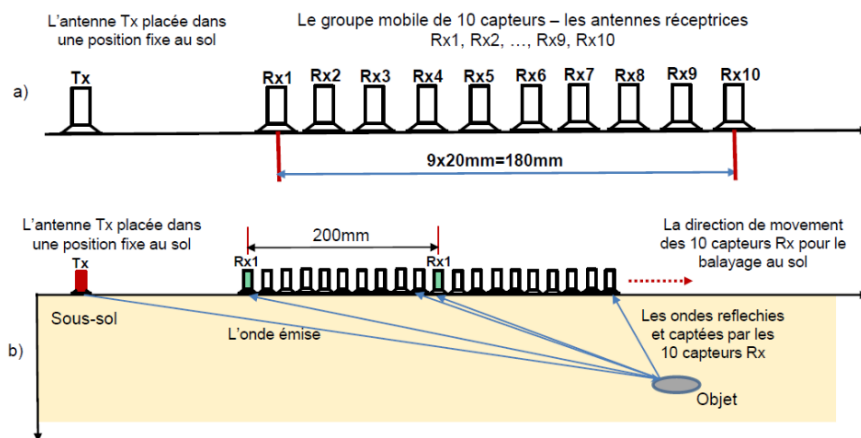


Figure 5.1: GPR multi-capteur

La figure 5.1 montre le principe du GPR multi-capteur utilisé pour l'acquisition des radargrammes. Comme vu plus haut, nous réalisons par ce procédé la capture de 10 A-scans par simulation. Les deux figures représentent:

- a) L'architecture du GPR multi-capteurs à 10 antennes réceptrices Rx et une antenne émettrice Tx au niveau du sol
- b) Le déplacement du groupe de 10 capteurs Rx par pas de 200 mm pendant le balayage d'exploration du sol.

5.2.2 Antennes Wu-King

Nous allons maintenant nous focaliser sur les antennes utilisées pour réaliser les simulations. Nous avons considéré des antennes de types Wu et King qui sont conçues pour fonctionner autour d'une fréquence de $f_0 = 1$ GHz. Cette d'antenne est de type dipôle amorti, elle se présente sous la forme d'un dipôle électrique présentant des résistances sur chacun des brins et un générateur de tension en son milieu. La forme temporelle de l'onde excitatrice au niveau du générateur est de type sino-gaussienne comme illustré Figure 5.2.

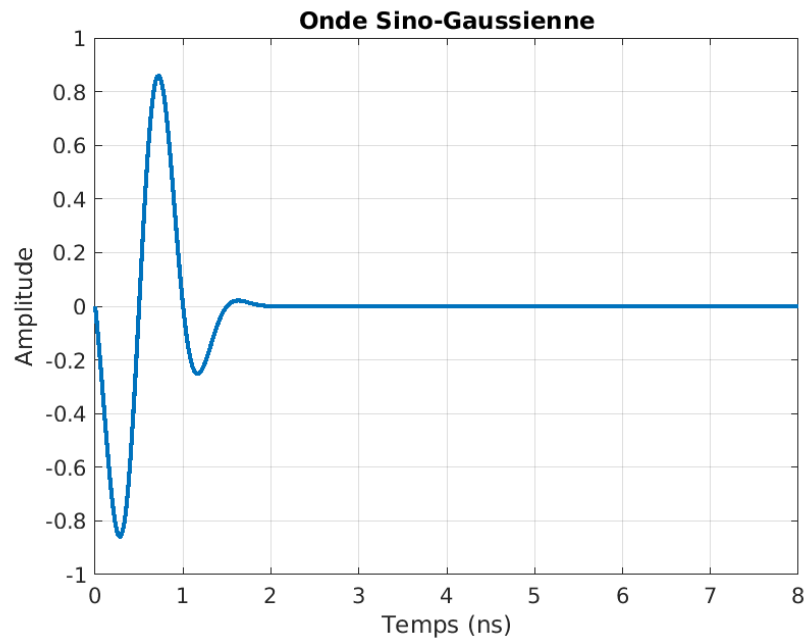


Figure 5.2: Réponse temporelle en courant avec un générateur de tension délivrant une forme d'onde sino-gaussienne

Le tableau 5.1 montre les différentes valeurs de résistance distribuées le long des brins utilisés dans toutes les simulation numériques effectuées au cours de cette thèse. Chacun brin est de taille de 25 cm, ce qui fait que la taille de l'antenne fait au total 50 cm. Sur chaque brin on crée 25 résistances (chacune d'entre elle répartie sur un segment FDTD de longueur 1 cm) de valeurs différentes, Au milieu, on connecte les deux brins à la source du signal.

Résistance	R_1	R_2	R_3	R_4	R_5	R_6	R_7	R_8	R_9
Valeurs (ohms)	39,0241	40.6847	42.4929	44,4694	46.6386	49.0303	51.6806	45.6338	57.9449

R_{10}	R_{11}	R_{12}	R_{13}	R_{14}	R_{15}	R_{16}	R_{17}	R_{18}	R_{19}
61.6833	65.9373	70.8126	76.4873	83.1384	91.0563	100.6412	112.4813	127.4788	147.091

R_{20}	R_{21}	R_{22}	R_{23}	R_{24}	R_{25}
173.8348	212.4647	273.1689	382.4365	637.3942	1912.1826

Table 5.1: Tableau des valeurs de résistance sur chaque brin pour une antenne de taille 50cm, R_1 représente la plus proche de la source d'alimentation, R_{25} correspond à la résistance la plus éloignée.

Le choix des valeurs des résistances données dans le tableau 5.1 plus haut suivant la loi décrite par Wu et King qui permet une adaptation large bande autour

de la fréquence de fonctionnement du dipôle non amorti. Ce choix est important pour assurer l'efficacité de l'antenne même si, il est bien connu qu'une grande partie de l'énergie est dissipée dans les résistances et seulement une faible partie est réellement rayonnée.

Le tableau 5.2 montre les valeurs choisies pour les paramètres de la simulation ainsi que les caractéristiques du radar ainsi constitué.

Paramètre de simulation	Valeur
Fréquence centrale de la source	1GHz
Forme d'onde de la source	impulsion sino-gaussienne
Durée de l'impulsion de la source	1ns
Intervalles entre les balayages A	20mm
Nombre de balayages A	40
Résolution spatiale	4mm
Pas de déplacement du réseau de capteurs	200mm

Table 5.2: Paramètres de simulation pour la simulation GPR.

5.2.3 Signal d'excitation

Pour exciter une scène donnée, il est important de choisir une excitation adaptée au problème traité. Nous allons préciser le choix de la fonction d'excitation dans le code électromagnétique.

Le logiciel TEMSI-FD dispose de plusieurs types de sources d'excitation, on peut citer:

- **Gaussienne:** la fonction source la plus courante car elle permet une réponse entre 0 et une fréquence maximale donnée.
- **Sino gaussienne:** une gaussienne centrée sur une certaine fréquence, permettant de créer un déphasage uniforme entre deux sources sur une large bande.

Sur toutes les simulations numériques, on utilisera un signal sino gaussienne.

La forme d'onde générée par l'antenne wu-king est une sino-gaussienne (sinus modulé par une gaussienne), cette impulsion est centrée à la fréquence $f_0 = 1$ GHz.

La forme temporelle et fréquentielle du signal sino-gaussien est représentée comme suit:

$$e(t) = A_0 \sin(2\pi f_0(t - t_0)) e^{-\frac{(t-t_0)^2}{T^2}} \xrightarrow{TFourier} e(f) = A_0 \frac{T\sqrt{\pi}}{2} e^{-j2\pi t_0 f} e^{-(\pi(f_0-f)T)^2} \quad (5.1)$$

- f_0 représente la fréquence centrale de la gaussienne.
- A_0 correspond à l'amplitude de l'impulsion de sino-gaussienne.
- t_0 représente le retard par rapport à l'origine des temps.
- T est la largeur de l'impulsion à mi-hauteur.

Trois approches sont envisageables pour établir les diverses valeurs de ces paramètres:

- Le domaine d'étude est une bande de fréquence connue, il suffit alors de la fixer et de donner l'atténuation $Att_{f_{max}}$ aux deux fréquences extrêmes de la bande $|f_0 - \Delta f, f_0 + \Delta f|$ par rapport au maximum d'amplitude situé à $f = f_0$
- Un paramètre indépendant est utilisé pour définir le pied de la gaussienne dans le domaine temporel. De ces données, les paramètres t_0 et T sont directement déterminés:

$$T = \frac{\sqrt{\ln(Att_{f_{max}})}}{\pi \Delta f} \quad t_0 = T \sqrt{\ln(Att_0)} \quad (5.2)$$

- La durée de la partie gaussienne de la fonction dans le domaine temporel est connue. Elle est égale à la valeur du temps t_{max} tel que:

On peut alors déduire les deux parametre t_0 et T :

$$e(t = t_{max}) = e(t = 0) = \frac{A_0}{Att_0}, t_0 = \frac{t_{max}}{2} \quad et \quad T = \frac{t_0}{\sqrt{\ln(Att_0)}} \quad (5.3)$$

Le temps de montée t_m de l'impulsion gaussienne doit être fixée alors les paramètres t_0 et T est défini comme suit:

$$T = \frac{t_m}{\sqrt{\ln(10)} - \sqrt{\ln\left(\frac{10}{9}\right)}} \quad et \quad t_0 = T \sqrt{\ln(Att_0)} \quad (5.4)$$

5.2.4 Milieu de propagation

Le milieu sur lequel vont être placées les antennes possède les caractéristiques données dans le tableau 5.3. Ainsi, quatre types de sol seront choisis pour générer la base de données.

Sl. No	Type de sol	Conductivité (S/m)	Permittivité relative (ϵ_r)
1	Dry sand	0.002	10.00
2	Dry clay	0.001	5.53
3	Concrete	0.001	8.00

Table 5.3: Paramètres diélectriques du type de sol utilisé dans la simulation.

Propriété	Valeur
eps_moyen	6
sige_moyen	1×10^{-6}
eps_graine	7
sige_graine	7
eps_facteur	1
sige_facteur	1.5×10^{-6}
eps_h	0.7
sige_h	0.7
eps_nit	4
sige_nit	3

Table 5.4: Propriétés du milieu fractal

Dans la scène, nous disposerons à chaque fois un tuyau situé à une profondeur choisie parmi 50 valeurs dans une plage variant entre $P_1 = 204$ mm et $P_{50} = 400$ mm, avec un pas de 4 mm. Les 50 profondeurs sont identifiées par leurs étiquettes $P_1, P_2, P_3, \dots, P_{50}$, la valeur d'une profondeur P_k étant calculée avec l'expression :

$$P_k = P_1 + (k - 1) \times 4 \text{ mm}, \quad k = 1, 2, 3, \dots, 50. \quad (5.5)$$

Un tuyau va être défini par sa position en profondeur P_k , son diamètre et sa constitution. Nous avons considéré 10 classes de diamètres présentées dans le tableau 5.5:

Index	1	2	3	4	5	6	7	8	9	10
Classe	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10
Valeur [mm]	16	24	32	40	48	56	64	72	80	100

Table 5.5: Les classes de diamètres pour la simulation des tuyaux.

Les tuyaux choisis peuvent aussi être de nature différente. La liste des types des tuyaux choisis pour cette simulation est donnée ci-après :

- Cylindre métallique complet,

- Cylindre rempli avec de l'eau,
- Cylindre creux et vide.

En résumé, nous nous proposons de réaliser une classification multiparamètres suivant:

- La nature du sol hôte,
- Le type de tuyau,
- Le diamètre des tuyaux.

5.3 Modélisation FDTD pour la création de la base de données

La base de données est générée à l'aide d'un solveur électromagnétique TEMSI-FD. Ce code est basé sur la méthode FDTD (Finite-Difference Time-Domain) et résout les équations vectorielles de Maxwell suivant le schéma numérique décrit dans chapitre 4. Ce simulateur est développé par l'équipe CEM du laboratoire XLIM de Limoges.

Dans le cadre de la modélisation de scènes GPR, il faut tout d'abord définir le volume de calcul, puis le milieu de propagation et enfin les antennes (émettrices et réceptrices) ainsi que les tuyaux enfouis.

Les dimensions de ce volume sont de 1 mètre de longueur, 0,8 mètre de largeur et 1,2 mètre de hauteur. Les mailles élémentaires sont cubiques, dans chaque direction, le pas delta fait 4 mm. Compte tenu du critère CFL, si les pas spatiaux sont égaux, le pas de temps est limité à [ref], ce qui fait $\Delta t_{\max} = 0.0077$ ns.

$$\Delta t_{\max} = \frac{1}{c} \cdot \frac{\Delta_{3D}}{\sqrt{3}} \quad (5.6)$$

La figure 5.3 illustre le modèle de la scène GPR avec un tuyau enfoui, l'antenne Tx localisée à sa position fixe à gauche de la scène et le réseau multi-capteurs positionné dans la première des 4 positions du balayage en direction x parallèle au sol. Le volume de calcul est défini comme suit :

- L'épaisseur de la couche de diélectrique représentant le sol est de 1.0 m, ses dimensions dans le plan horizontal sont infinies (plongées dans les PML sur les côtés),
- Les 0.2 m supérieurs représentent l'air
- Chaque antenne est un dipôle électrique de Wu-King de 0.5 m de long. En

vert sur la figure, nous distinguons l'antenne émettrice isolée et le réseau de 10 antennes réceptrices qui va se déplacer suivant la direction x.

- En rouge est représenté le tuyau enfoui de diamètre et de constitution variables comme défini plus haut. Ce tuyau est supposé infini suivant l'axe y.

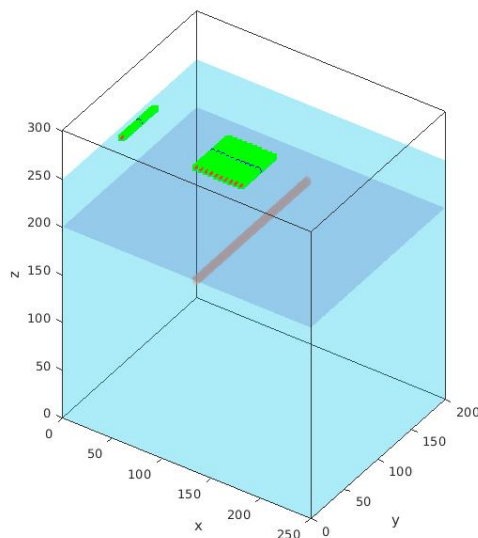


Figure 5.3: Scène GPR avec un tuyau métallique de diamètre 32mm enfoui dans un sol d'argile sèche

5.4 Tracé des B-scans et recalage temporel

La localisation précise d'un objet à l'aide d'un seul A-scan est impossible. Un A-scan indique uniquement le temps écoulé depuis l'émission de l'onde par la source, la trajectoire parcourue dans le sol entre différents objets et l'antenne de réception. En fait, il faudrait être exactement au-dessus du tuyau et connaître les caractéristiques du milieu pour avoir la vitesse de propagation et ainsi déduire la profondeur du tuyau.

Aussi, suite à cette remarque, on voit qu'il est nécessaire d'effectuer plusieurs mesures à des positions en x différentes pour connaître cette position. Aussi, l'analyse est classiquement faite en considérant plusieurs A-scans. Le tracé classique consiste à représenter ceux-ci de la manière suivante :

- Pour chaque position du radar, on trace la variation temporelle du signal reçu

sur un axe vertical,

- On déplace le radar et on positionne le tracé suivant à une cote x décalée du pas de déplacement.

Au final, on obtient une image du cylindre appelée B-scan ou radargramme qui se présente sous la forme d'une hyperbole dont :

- Le haut de l'hyperbole correspond à la cote x de positionnement du tuyau,
- Les branches de l'hyperbole ont des pentes liées à la vitesse de propagation dans le milieu. Ce qui permet d'avoir une première idée de la permittivité de ce dernier.

La qualité des images 2D B-scan est considérablement améliorée en éliminant les échos (réflexions) provenant de la surface du sol ainsi que les couplages directs entre les antennes. Une deuxième simulation de la même scène sans objet est réalisée pour capturer une image B-scan ne contenant que les échos (réflexions) de la surface du sol et les couplages directs. D'un point de vue pratique, en configuration monostatique, le couplage entre antennes est toujours identique pour un sol supposé homogène, de ce fait il est invariant lorsque le radar se déplace et donne une bande horizontale sur le radargramme qu'il est possible de soustraire de l'image. Dans notre configuration, le problème est plus complexe d'un point de vue expérimental, et il est nécessaire de calibrer le signal de couplage en fonction de la distance entre antenne d'émission et de réception. Toutefois, dans le cas présent qui concerne une problématique uniquement théorique, il sera nécessaire de réaliser pour toutes les positions du radar une simulation avec le tuyau et une sans le tuyau. Nous pouvons toutefois remarquer que, pour un sol à caractéristiques fixé, une seule simulation pourra être utilisée pour tous les types et diamètres de tuyaux. Ainsi, en soustrayant l'image B-scan de la deuxième simulation de l'image B-scan obtenue à partir de la première simulation, on obtient une image B-scan où seules les hyperboles dues aux réflexions provoquées par les objets enfouis sont visibles.

La Figure 5.4 présente la création des images B-scan où les échos de la surface du sol et les couplages directs entre les antennes ont été éliminés en soustrayant l'image B-scan de la deuxième simulation

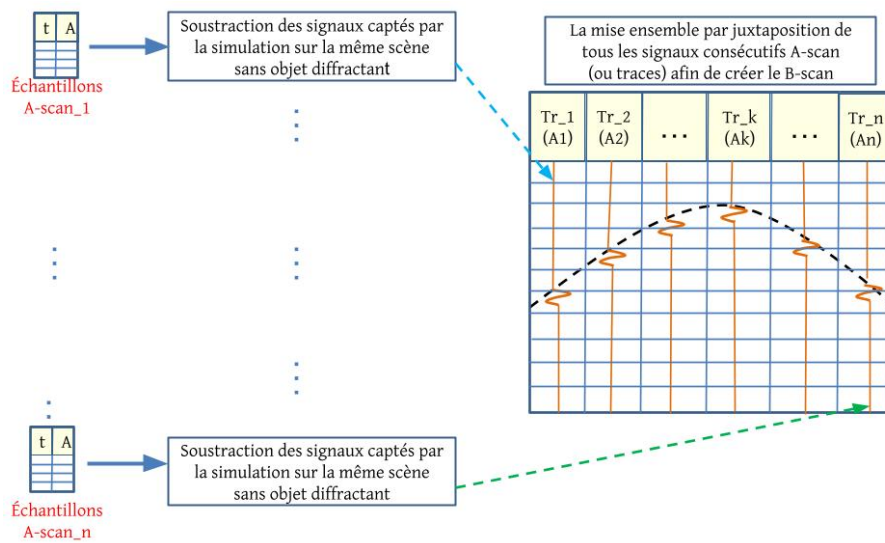


Figure 5.4: La création de l'image B-scan par la juxtaposition de tous les signaux consécutifs A-scan, après soustraction, pour chaque A-scan, des signaux captés sur la même scène sans objet diffractant

Lors de la simulation, chaque antenne réceptrice du groupe des 10 capteurs enregistre un signal A-scan. A chaque position de balayage du groupe des 10 capteurs, les enregistrements A-scan ont des déphasages (décalages temporels) différents qui sont déterminés par la position du groupe et donc des temps de propagation différents entre l'objet enfoui et les antennes réceptrices.

Considérons les différentes positions des antennes qui sont au nombre de $N \times 10$, N étant le nombre de positions du réseau ; rappelons que l'émetteur reste fixe. Le schéma de la Figure montre le retard entre l'impulsion émise et celle reçue par l'antenne réceptrice provenant de la diffraction du tuyau. Dans ce cas, le retard est la somme des deux temps de parcours : émetteur - tuyau et tuyau - récepteur. Lorsque l'on se trouve au sommet de l'hyperbole, l'antenne réceptrice passe au-dessus de l'emplacement du tuyau. Si l'on connaît la permittivité relative du milieu, il est alors possible de traduire les retards en distance.

Dans ce cas précis, on connaît la distance entre l'émetteur (T_x) et le récepteur (R_x), d'une part $dx = x_k - x_{T_x}$, et la distance totale parcourue donnée par le retard de propagation, d'autre part dT . En appliquant le théorème de Pythagore, on peut déduire la profondeur du tuyau (dP) et ainsi sa position en x et z .

La figure 5.5 montre l'enregistrement des signaux A-scan par un GPR multi-capteur; a) la position fixe de l'antenne source T_x et la position variable au sol x_k d'une antenne réceptrice R_{xn} du GPR multi-capteur; b) la position temporelle de l'échantillon du signal A-scan sur l'hyperbole, déterminée par le temps de prop-

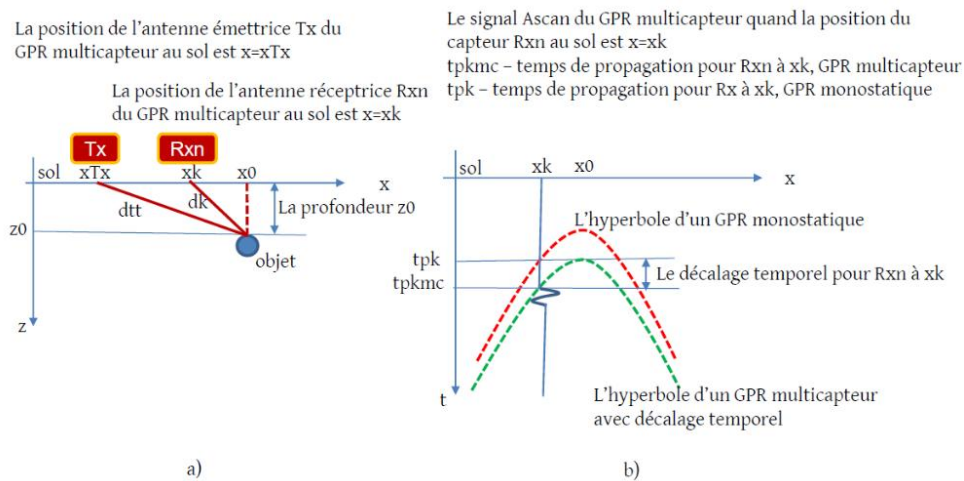


Figure 5.5: Schéma illustrant le retard entre l'impulsion émise et reçue due à la diffraction du tuyau.

agation tp_{kmc} qui montre le décalage temporel par rapport à un enregistrement monostatique à la position x_k dont le temps de propagation est $tp_k < tp_{kmc}$. Ainsi, on connaît le temps de propagation entre l'émetteur et le tuyau, ce temps pourra être soustrait de façon systématique sur tous les A-scan. En divisant par deux les temps (ou distances) obtenus, on obtiendra directement le temps de propagation entre le tuyau et n'importe quelle position du récepteur.

5.5 Prétraitement des données

Dans cette section, nous aborderons le prétraitement des données, une étape cruciale dans la mise en œuvre des modèles d'apprentissage profond en général et dans le contexte du radar à pénétration de sol (GPR) en particulier. Le processus de traitement des données GPR dans cette étude comporte plusieurs étapes essentielles, notamment l'extraction de caractéristiques pertinentes, la segmentation du jeu de données en ensembles d'apprentissage et de test, l'entraînement de modèles de réseaux de neurones convolutifs (CNN) et la validation des résultats obtenus. Parmi les opérations de traitement, nous utilisons des méthodes telles que Sobel, Prewitt et Canny, entre autres, pour la détection de contours. Cette détection de contours revêt une importance cruciale en traitement d'image et en vision par ordinateur, car elle permet de délimiter précisément les objets présents dans une image. L'architecture du CNN proposée est évaluée dans le contexte du traitement d'images, et ses performances sont comparées en utilisant divers

opérateurs, notamment Sobel, Laplacian, Scharr et Canny. L'efficacité du modèle du CNN est un point d'intérêt majeur. Les images GPR se caractérisent par des signatures hyperboliques qui les distinguent des autres types d'images. Ces signatures hyperboliques sont composées de paramètres permettant d'estimer la taille, l'emplacement de l'objet ciblé, ainsi que d'autres caractéristiques de l'image. Cette estimation inclut également la reconnaissance des propriétés environnementales entourant l'objet, comme le sol. Lors de la détection des bords, l'objectif est de repérer les zones d'une image présentant un changement brusque de couleur ou d'intensité. Ce changement soudain est directement proportionnel à la valeur numérique mesurée : plus cette valeur est élevée, plus le changement est perceptible, tandis qu'une valeur plus basse indique un changement moins marqué.

5.5.1 Labellisation des images B-scan

Pour l'étiquetage efficace des images radar B-scan, nous avons développé une fonction spécifique qui tire parti du nom du fichier pour générer les étiquettes associées à chaque image. La fonction d'étiquetage commence par diviser le nom du fichier en sections en utilisant le caractère '.' comme séparateur, créant ainsi la liste 'name_parts'. Par la suite, la première partie de cette liste est divisée en utilisant le caractère '_' comme séparateur, ce qui nous permet d'obtenir une liste de labels que nous stockons dans une variable appelée 'labels'. Par exemple, si le nom de fichier est 'diam100_argile_eau1.png', la fonction générerait trois étiquettes distinctes : 'diam100', 'argile' et 'eau'. Ensuite, des opérations de traitement supplémentaires sont appliquées aux étiquettes extraites, telles que la suppression de chiffres indésirables lorsque cela est nécessaire. Ces étiquettes fournissent des informations essentielles concernant les caractéristiques géométriques, la composition du milieu, ainsi que le type de tuyau représenté dans chaque image radar B-scan. Nous sommes ainsi en mesure d'automatiser le processus d'étiquetage des images radar B-scan, ce qui simplifie grandement la classification et l'analyse ultérieure des données. En résumé, l'étiquetage des images consiste à attribuer des informations descriptives à chaque image pour faciliter leur compréhension et leur utilisation dans des applications telles que la classification et l'analyse de données.

Quelques exemples de radargrammes sont donnés ci dessous:

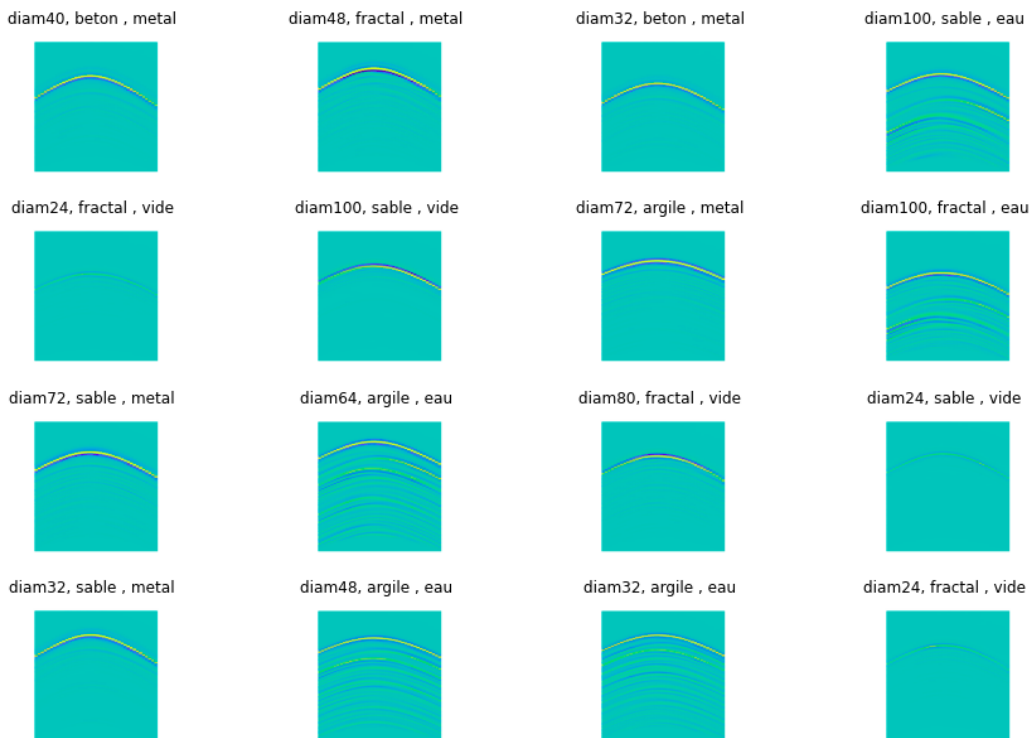


Figure 5.6: Exemples de radargrammes.

5.6 Détection des contours

5.6.1 Principe d'opérateur de gradient

La technique de l'opérateur de gradient est largement utilisée dans le traitement d'image pour reconnaître les bords des objets présents dans une scène. Le principe fondamental de cette technique s'articule autour de l'examen des changements d'intensité des pixels d'une image afin de détecter les transitions nettes qui correspondent aux contours. Cela permet de détourer les objets avec une grande précision et exactitude. Pour simplifier le processus de détection des contours, les images initialement en couleur sont souvent converties en niveaux de gris. Cette conversion permet de se concentrer sur les variations d'intensité plutôt que sur la couleur, simplifiant finalement le processus. Pour calculer les gradients d'une image, des opérateurs différentiels comme Sobel, Prewitt ou Canny sont utilisés. Ces opérateurs calculent les dérivées partielles de l'intensité de l'image le long des axes horizontal et vertical ils permettent de mesurer les fluctuations d'intensité entre pixels adjacents. Pour déterminer la taille et l'orientation du dégradé, les différentes étapes sont décrites ci-après. Tout d'abord, les gradients horizontaux

et verticaux sont mesurés puis additionnés pour déterminer l'amplitude du gradient. L'amplitude est une mesure de l'intensité du gradient à chaque pixel, ce qui correspond à l'intensité du bord de l'objet. De plus, la direction du gradient peut être calculée, fournissant la direction dans laquelle le décalage d'intensité maximum se produit. Le processus de détermination des pixels appartenant aux bords est appelé seuillage. Afin d'identifier ces pixels, l'amplitude du gradient est évaluée et comparée à une valeur de seuil prédéterminée. Si l'amplitude du gradient dépasse ce seuil, le pixel est considéré comme faisant partie d'un bord. Inversement, si l'amplitude du gradient tombe en dessous du seuil, le pixel est considéré comme faisant partie de l'arrière-plan. Le seuil est défini comme la valeur minimale requise pour qu'un pixel soit reconnu comme un bord. Nous allons décrire plus en détail les principaux filtres classiquement utilisés.

5.6.2 Sobel

L'utilisation de l'opérateur de Sobel suit un algorithme simple de détection des contours d'une image. Il permet le calcul de la dérivée première de la fonction "niveau de gris" des points de l'image. Cet algorithme recherche les instances de variation maximale du niveau de gris. En utilisant deux masques de convolution de 3x3 pixels; l'opérateur les applique à l'image d'origine, ce qui donne une carte de dégradé de niveaux de gris. Les points lumineux sur cette carte indiquent les zones où il y a un changement rapide de l'intensité du niveau de gris sur une courte plage, indiquant ainsi les contours de l'image d'origine.

Pour calculer la dérivée dans la direction X, l'image $f(x, y)$ est convoluée avec un noyau SobelX. Nous nous concentrerons sur une taille de noyau de 3x3.

Pour obtenir une approximation de la norme du gradient, les approximations du gradient vertical SobelY et horizontal SobelX peuvent être combinées en chaque point, de même, on peut calculer la direction du gradient de la manière suivante:

$$SobelX = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \quad (5.7)$$

$$SobelY = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} \quad (5.8)$$

Pour calculer l'amplitude du gradient (approximation de la norme) en chaque point de l'image:

$$\text{Amplitude du gradient} = \sqrt{\text{Sobel}X^2 + \text{Sobel}Y^2} \quad (5.9)$$

Pour calculer la direction du gradient en chaque point de l'image (en radians):

$$\text{Direction du gradient} = \arctan\left(\frac{\text{Sobel}Y}{\text{Sobel}X}\right) \quad (5.10)$$

5.6.3 Canny

L'opérateur de gradient de Canny, développé par John Canny en 1986, est une référence dans le traitement d'images pour la détection des bords. Il s'appuie sur une méthodologie précise et rigoureuse pour obtenir des résultats d'une grande finesse. Sa précision et sa robustesse proviennent de plusieurs étapes qui garantissent une détection de bords avec une grande précision.

- **Réduction du bruit:** Au début du processus, un filtre gaussien est utilisé pour éliminer le bruit de l'image. Cela sert à rendre l'image homogène, prévenant ainsi toute variation indésirable d'intensité qui pourrait causer des artefacts.
- **Calcul du gradient:** L'étape suivante consiste à déterminer la direction et la magnitude des changements d'intensité dans l'image. Pour cela, des opérateurs de dérivée partielle, comme le filtre de Sobel, sont couramment utilisés. Chaque pixel de l'image est analysé pour ses gradients horizontaux et verticaux, fournissant des informations précises sur le changement d'intensité.
- **Suppression des non-maxima:** Une fois le gradient calculé, une étape cruciale consiste à raffiner la précision des bords. L'idée est d'éliminer les pixels qui n'ajoutent pas de valeur significative à la détection des bords. Cette étape assure que seuls les bords les plus pertinents sont mis en avant.
- **Double seuillage:** La détection est ensuite renforcée par l'utilisation de deux seuils, un élevé et un bas. Les bords avec une amplitude de gradient supérieure au seuil élevé sont conservés, tandis que ceux en dessous du seuil bas sont éliminés. Les bords intermédiaires, dont l'amplitude se situe entre les deux seuils, sont conservés s'ils sont connectés à des bords forts.
- **Suivi de contour par hystérésis:** La dernière étape du processus renforce davantage les bords faibles probablement liés aux bords forts. Grâce à la connectivité des pixels, les bords faibles associés aux bords forts sont

renforcés, assurant ainsi une détection de bord optimale.

Dans ce cas, nous avons utilisé l'opérateur de gradient de Canny à travers la fonction OpenCV `cv2.Canny()`. La détection des bords est affichée à travers plusieurs images:

- Original Image: L'image initiale sans modifications.
- Canny Edges: Les bords détectés après avoir appliqué l'opérateur de Canny.
- Bounding Box: L'image originale avec une boîte englobante mise en évidence en rouge, délimitant la zone contenant les bords les plus significatifs.
- Cropped: Une image recadrée et redimensionnée basée sur la boîte englobante précédemment définie.

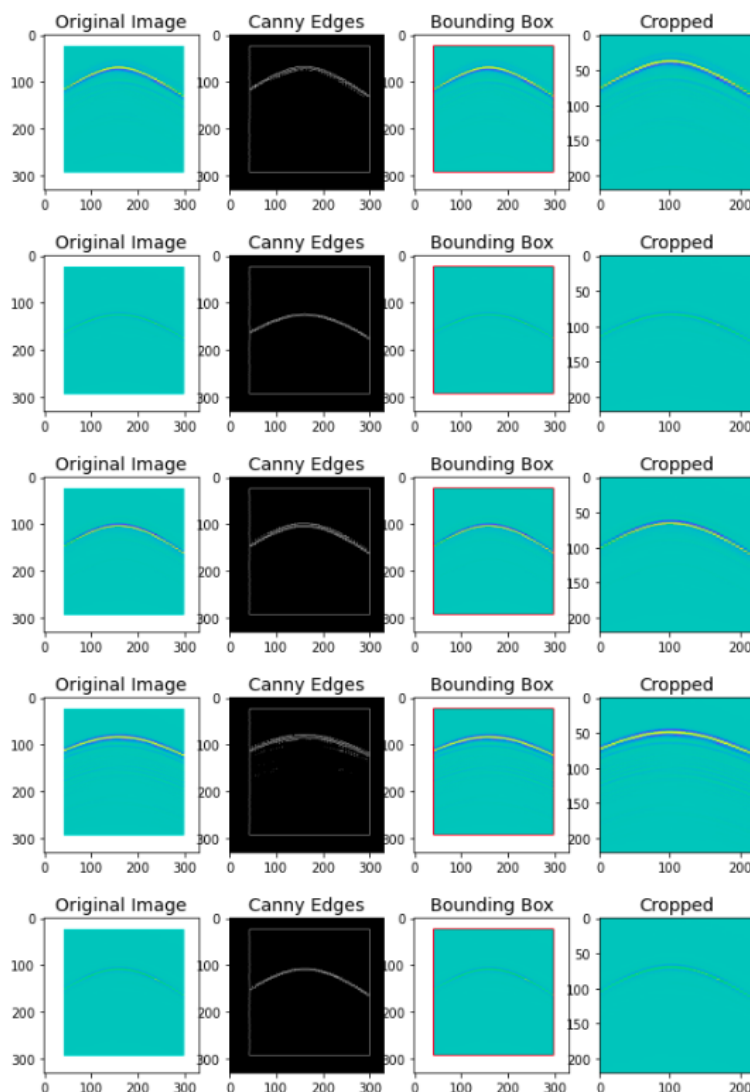


Figure 5.7

5.6.4 Augmentation des données (data augmentation)

Dans cette partie, nous allons utiliser la classe `ImageDataGenerator` de la bibliothèque Keras/Tensorflow pour augmenter notre base de données. Cette technique consiste à augmenter la taille de l'ensemble de données d'entraînement en appliquant des transformations aléatoires aux images existantes. Cela permet d'améliorer la capacité du modèle à généraliser et à traiter différentes variations des images. Il s'agit par exemple de décalage horizontal des hyperboles dans la scène; en effet, dans nos simulations les positions des tuyaux (et donc les hyperboles) sont tous centrés dans l'image ce qui peut être réducteur suivant les zones investiguées dans la réalité. Ces transformations contribuent à diversifier notre ensemble de données et à rendre le modèle plus robuste face à différentes conditions et variations des images.

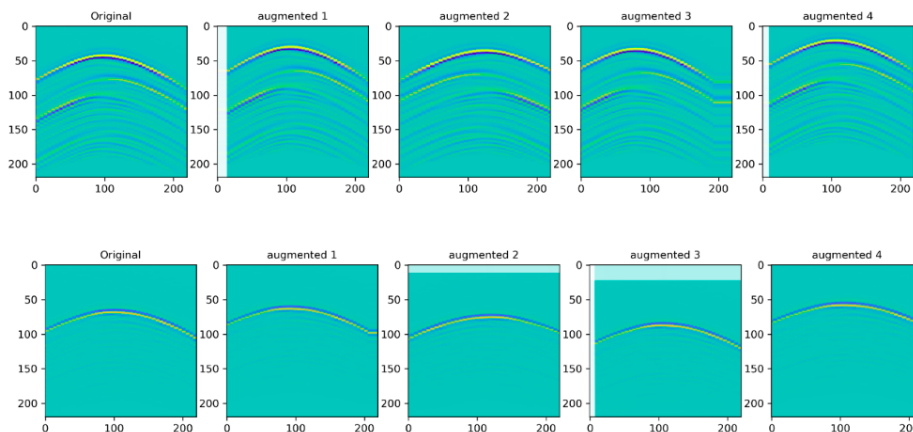


Figure 5.8: Exemples d'augmentation des données radagrammes

5.6.5 Séparation des données

Dans chaque modèle d'apprentissage, il est primordial de diviser les données en différents sous-ensembles : d'apprentissage, de validation, et de test. Cette séparation est cruciale pour évaluer et optimiser les performances du modèle. L'évaluation des performances se fait en répartissant les données en un ensemble d'apprentissage (généralement 60%), un ensemble de validation (généralement 20%) et un ensemble de test (généralement 20%). Il est à noter que ces trois sous-ensembles sont distincts et disjoints. Alors que les données d'apprentissage servent à former le modèle et celles de validation à optimiser ses hyperparamètres, l'ensemble de test, lui, est réservé pour évaluer l'exactitude, la précision, le rappel

et d'autres indicateurs de performance du modèle sur des données inédites. Cela offre une vision de la capacité du modèle à se généraliser face à de nouvelles données.

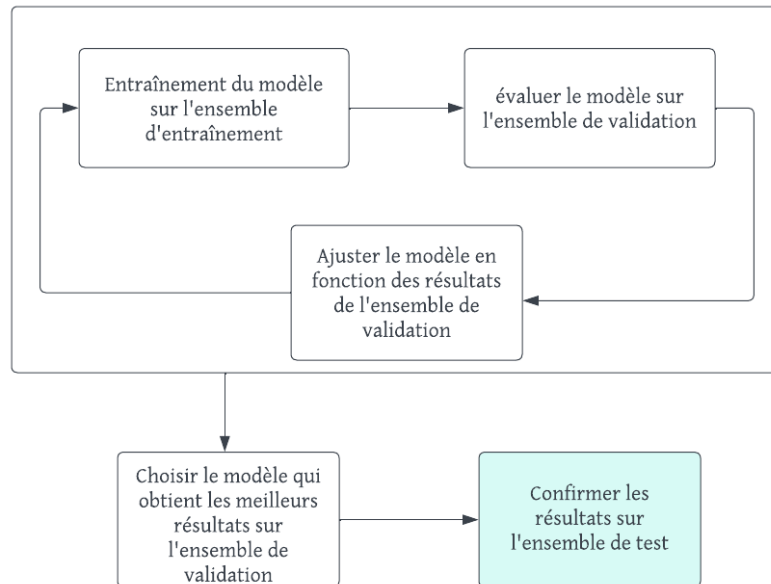


Figure 5.9: Illustrations de schémas de partitionnement des données d'entraînement, validation et de test.

5.6.6 Encodage des étiquettes

Pour réaliser l'opération d'encodage des sorties pour la classification, on utilise une technique couramment appelée "encodage one-hot", qui consiste à une représentation de chaque variable catégorielle sous forme d'un vecteur binaire. La stratégie consiste à créer des data frame que intègrent les différents labels, puis en appliquant l'encodage one-hot aux colonnes spécifiées, on génère un nouveau Data Frame qui contient les données transformées. Cette opération permet de représenter les labels catégoriques sous forme de vecteurs binaires pour une utilisation ultérieure dans des modèles d'apprentissage profond. Chaque label est représenté comme un vecteur binaire où chaque position dans le vecteur représente un label potentiel. Si le label est présent, alors la valeur à cette position est 1, sinon c'est 0.

Pour l'exemple donné avec trois labels possibles : "vide", "argile", et "métal"

Index	diam100	argile	métal
0	1	0	0
1	0	1	0
2	0	0	1
3	1	1	0
4	0	1	1
5	1	0	1

Table 5.6: Exemple d'encodage des étiquettes

5.7 Architectures utilisées pour la multi label classification des images B-Scans

5.7.1 Principe de transfert par apprentissage (transfert Learning)

Le principe de transfert d'apprentissage est une technique utilisée dans l'apprentissage profond, qui consiste à utiliser les connaissances et les modèles acquis à partir d'une tâche source pour résoudre une autre tâche liée. Cette approche élimine le besoin de créer un modèle entièrement nouveau pour la tâche cible. Pour ce faire, il exploite les informations et les représentations obtenues à partir de la tâche source permettant d'améliorer les performances lors de la résolution de la tâche cible. Après avoir prétraité les données et encodées les classes qui seront les sorties de notre réseau de neurones, nous allons procéder à la construction du réseau . Pour ce faire, nous avons testé trois types de réseau : Resnet 50, VGG et un réseau personnalisé. Dans un premier temps, nous allons décrire la constitution de ces réseaux , puis évaluer les performances sur l'étude des radargrammes par la suite.

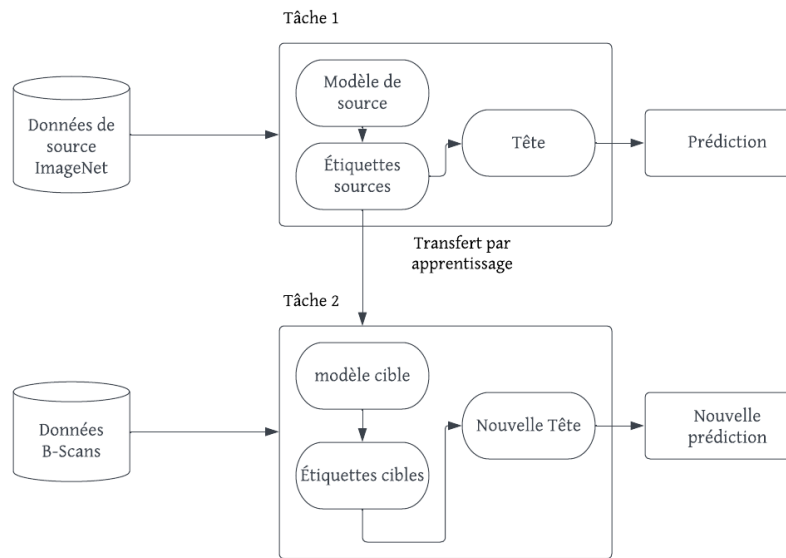


Figure 5.10: Stratégie globale de transfert par apprentissage.

5.7.2 Architecture Resnet 50

Pour établir notre modèle de réseaux de neurones, nous proposons dans un premier temps d'utiliser une architecture de type ResNet-50, ce modèle sera terminé par 4 couches spécifiques. La première opération pour utiliser le processus d'apprentissage par transfert learning est de charger les poids du réseau qui ont été préalablement entraînés sur la base de données ImageNet. La taille de l'entrée du ResNet-50 est de $220 \times 220 \times 3$, les couches de pooling reposent sur l'utilisation d'un pooling moyen. La Figure 5.11 présente l'architecture du ResNet-50 adaptée pour le transfert d'apprentissage. Dans le cadre de cette thèse, le modèle (ResNet-50 pré-entraîné) a été adapté à notre problématique et le transfert d'apprentissage a été appliqué pour classifier notre base de données d'images B-Scans du radar GPR afin de réaliser une classification Multi-label. L'architecture est composée de 50 couches. La première est une couche de convolution qui est suivie par une couche de pooling, comme vu dans le chapitre 1, l'objectif de cette couche est de réduire la résolution spatiale de nos B-Scans. En réduisant la résolution de nos radargramme, le nombre de paramètres diminue. ce processus simplifie le modèle et diminue les besoins en termes de calculs. Ensuite, L'architecture comporte une série de blocs résiduels qui seront empilés les uns sur les autres. Chaque bloc résiduels contient une série de couche de

convolutions avec des opérations qui permettent aux réseaux d'apprendre et de capturer des caractéristiques les plus complexes (formes et contours d'objets, identification du milieu de propagation,...). L'architecture Resnet-50 est composée de 16 blocs résiduels. La dernière couche FC (Fully Connected) du modèle utilise une couche de pooling moyen : elle calcule la moyenne des valeurs dans chaque canal de la carte des caractéristiques et une couche entièrement connectée. Les couches supplémentaires remplacent la couche entièrement connectée du ResNet-50. Cette couche représente trois branches de sortie pour notre modèle ResNet. Chaque branche est construite à partir de la sortie de la dernière couche de ResNet avec des couches entièrement connectées supplémentaires et une couche de sortie correspondant au nombre de classes de chaque tâche de classification spécifique. Cela permet au modèle de prédire l'ensemble des étiquettes de façon simultanément.

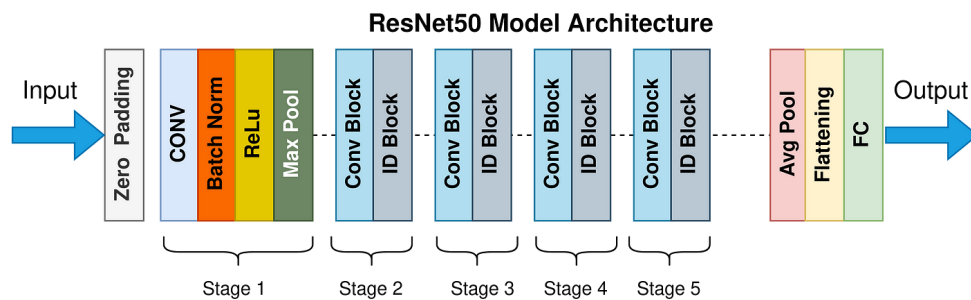


Figure 5.11: Architecture Resnet-50

Branche A (geometry_out):

$$\begin{aligned}
 a_1 &= \text{Flatten}(O_{\text{resnet}}) \\
 a_2 &= \sigma(W_{a1} a_1 + b_{a1}) \\
 a_3 &= \sigma(W_{a2} a_2 + b_{a2}) \\
 z_a &= W_{a3} a_3 + b_{a3} \\
 O_{a,i} &= \frac{e^{z_{a,i}}}{\sum_{j=1}^9 e^{z_{a,j}}}
 \end{aligned}$$

Branche B (medium_out):

$$b_1 = \text{Flatten}(O_{\text{resnet}})$$

$$b_2 = \sigma(W_{b1}b_1 + b_{b1})$$

$$b_3 = \sigma(W_{b2}b_2 + b_{b2})$$

$$z_b = W_{b3}b_3 + b_{b3}$$

$$O_{b,i} = \frac{e^{z_{b,i}}}{\sum_{j=1}^4 e^{z_{b,j}}}$$

Branche C (type_out):

$$c_1 = \text{Flatten}(O_{\text{resnet}})$$

$$c_2 = \sigma(W_{c1}c_1 + b_{c1})$$

$$c_3 = \sigma(W_{c2}c_2 + b_{c2})$$

$$z_c = W_{c3}c_3 + b_{c3}$$

$$O_{c,i} = \frac{e^{z_{c,i}}}{\sum_{j=1}^3 e^{z_{c,j}}}$$

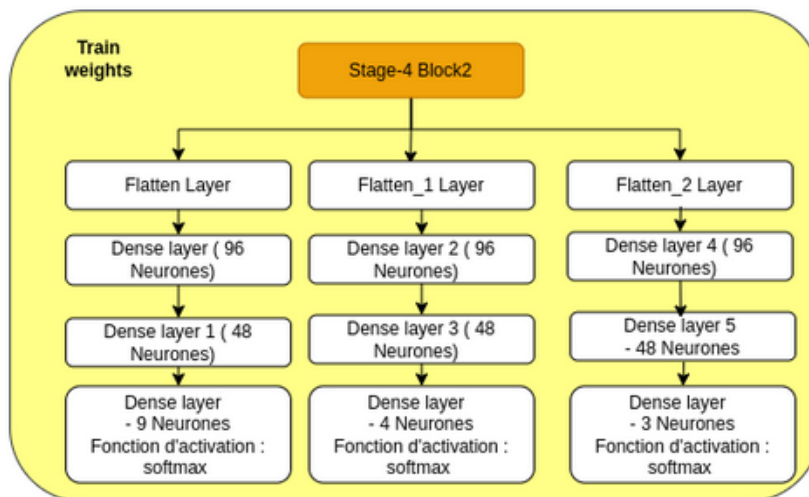


Figure 5.12: Multi-branches entièrement connectées

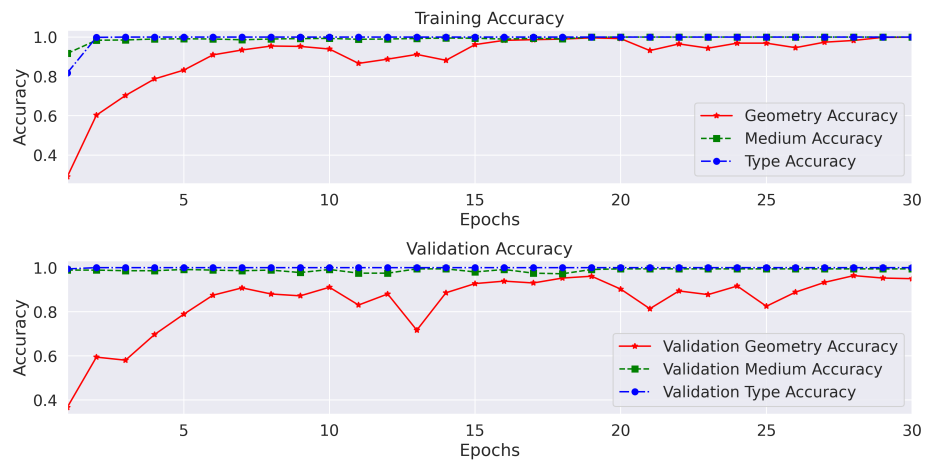


Figure 5.13: Exactitude de la validation pour ResNet-50 au fil des époques.

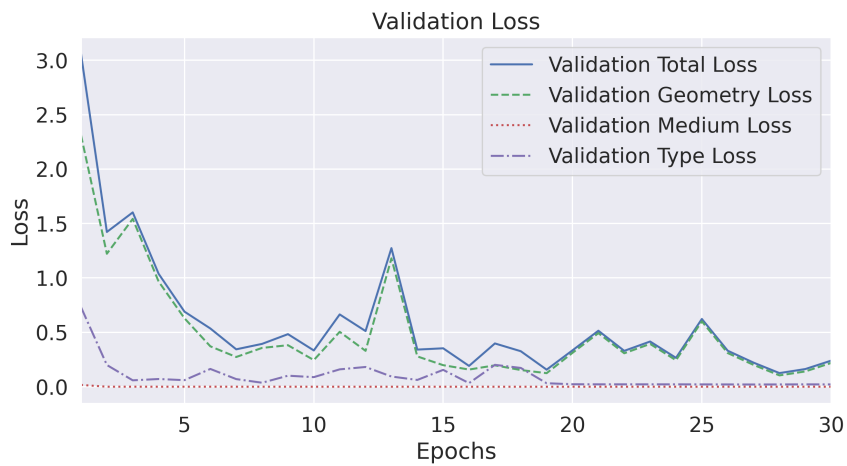


Figure 5.14: Perte de la validation pour ResNet-50 au fil des époques.

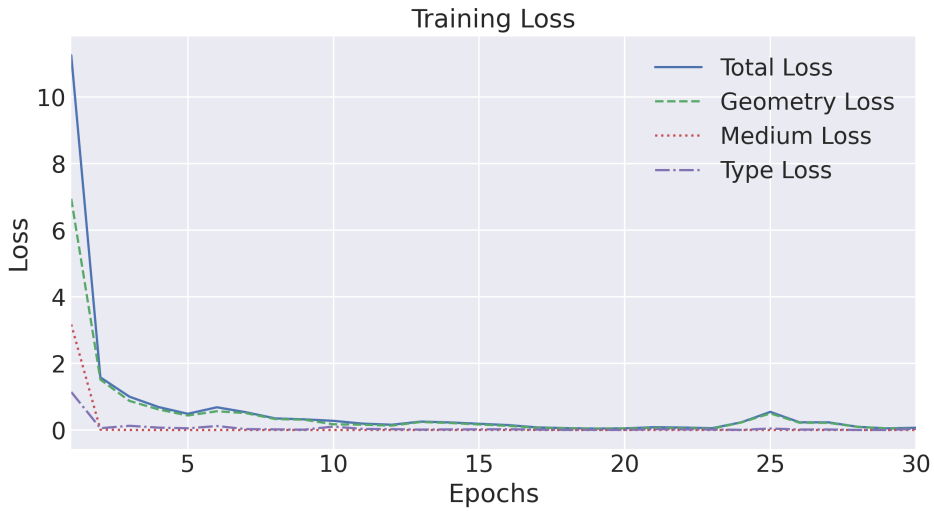
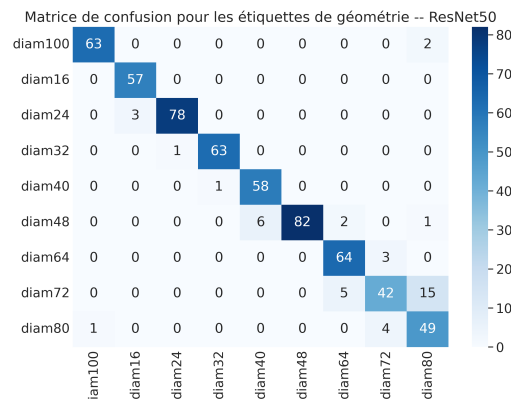
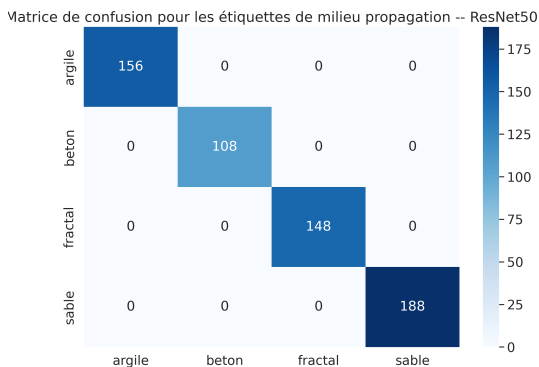


Figure 5.15: Perte d’entraînement pour ResNet-50 au fil des époques.

Les figure 5.13, 5.14, et 5.15 montrent respectivement les diagrammes de pertes, de précision de l’entraînement et de validation. On constate que, pour la première époque, la perte et la précision s’améliorent de manière exponentielle, puis présentent une allure quasi-horizontale entre les époques 20 et 30. Au cours de cette période, une certaine instabilité de l’apprentissage a été observée et, vers la fin, une amélioration significative a été constatée. Grâce à l’utilisation d’une technique d’arrêt précoce, l’entraînement s’est arrêté à la 30e époque (sur un total de 40 époques). Les graphiques de pertes et d’exactitude ont montré un apprentissage régulier dès la première époque (Figure x) sur l’ensemble des trois classes [géométrie, type de tuyau ainsi que le milieu de propagation].



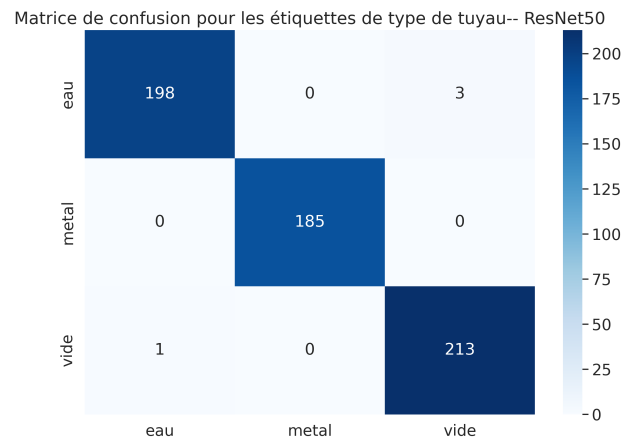


Figure 5.16: Matrice de confusion pour l'architecture ResNet-50.

La figure 5.16 illustre la matrice de confusion qui présente les résultats d'une classification impliquant neuf classes de diamètres différents (diam16, diam24, diam32, diam40, diam48, diam64, diam72, diam100), ainsi que les résultats de quatre types de tuyaux et les résultats pour chaque milieu de propagation. Concernant la matrice de confusion pour la classification du 'type de tuyau', les performances du modèle pour trois classes - eau, métal et vide - sont illustrées. Le modèle a réalisé des prédictions à environ 98,51% de justesse pour la classe "eau", avec 3 prédictions incorrectes confondues avec la classe "vide". Pour la classe "métal", toutes les 185 prédictions étaient correctes, atteignant ainsi une précision de 100%, sans aucune erreur. En ce qui concerne la classe "vide", le modèle a produit des prédictions avec environ 99,53% de justesse, avec une unique prédiction incorrecte confondue avec la classe "eau".

$$\text{Précision} = \left(\frac{\text{Nombre de Prédictions Correctes}}{\text{Nombre Total de Prédictions pour la Classe}} \right) \times 100$$

Pour la classe "eau":

$$\text{Précision} = \left(\frac{198}{198 + 3} \right) \times 100 \approx 98.51\%$$

Pour la classe "métal":

$$\text{Précision} = \left(\frac{185}{185} \right) \times 100 = 100\%$$

Pour la classe "vide":

$$\text{Précision} = \left(\frac{213}{213 + 1} \right) \times 100 \approx 99.53\%$$

Pour la matrice de confusion associée à la classe du milieu de propagation (argile, béton, fractal et sable), le modèle a affiché d'excellentes performances. Il a réalisé des prédictions correctes pour chaque classe : respectivement 156 pour l'argile, 108 pour le béton, 148 pour le fractal et 188 pour le sable, sans aucune erreur.

Concernant la matrice de confusion pour la classe de diamètre, le modèle a également montré de bonnes performances globales, la majorité des observations étant correctement classées pour la plupart des classes de diamètre. Toutefois, des confusions ont été observées entre les classes de diamètres 24mm, 48mm, 64mm et 72mm. Ces erreurs pourraient être attribuées à des similarités dans les caractéristiques des différentes classes.

En résumé, le modèle possède de bonnes performances globales, la plupart des prédictions sont correctes pour toutes les classes. Cependant, des confusions sont observées entre certaines classes, en particulier pour les classes de diamètre spécifiques.

Rapport de classification	Précision	Recall	F1-Score
diamètre 16 mm	0.95	1.00	0.97
diamètre 24 mm	0.99	0.96	0.97
diamètre 32 mm	0.98	0.98	0.98
diamètre 40 mm	0.91	0.98	0.94
diamètre 48 mm	1.00	0.90	0.95
diamètre 64 mm	0.90	0.96	0.93
diamètre 72 mm	0.86	0.68	0.76
diamètre 80 mm	0.73	0.91	0.81
diamètre 100 mm	0.98	0.97	0.98
Macro average	0.92	0.93	0.92
Weighted average	0.93	0.93	0.93

Table 5.7: Rapport de classification des classes de diamètre via le modèle ResNet-50

Le tableau 5.7, représente le rapport de classification qui présente les mesures de précision, de rappel et de score F1 pour chaque classe de la catégorie "Geometry" Sur l'ensemble de test. Pour la classe "diamètre 100", la précision est de 0,98, ce qui signifie que 98 % des prédictions pour cette classe sont correctes. Le rappel est de 0,97, ce qui indique que 97 % des instances réelles de cette classe ont été correctement identifiées. Le score F1 est de 0,98, ce qui

représente une mesure globale de la performance du modèle pour cette classe. Pour la classe "diametre 16", le modèle a obtenu une précision de 0,95, un rappel de 1,00 et un score F1 de 0,97. Cela indique que les prédictions pour cette classe sont généralement précises, avec un rappel de 100 %, ce qui signifie que toutes les instances réelles de cette classe ont été correctement identifiées. Pour les autres classes de diamètres (24, 32, 40, 48, 64, 72, et 80), le modèle a obtenu des mesures de précision, de rappel et de score F1 variant entre 0,90 et 0,99. Ces mesures indiquent des performances globalement bonnes pour ces classes, bien que certaines classes présentent des scores F1 légèrement inférieurs, suggérant une certaine difficulté dans la prédiction précise de ces classes. En ce qui concerne les mesures agrégées, la précision moyenne (macro avg) est de 0,92, le rappel moyen est de 0,93 et le score F1 moyen est de 0,92. Cela montre une très bonne performance globale du modèle sur l'ensemble des classes. En conclusion, le modèle présente des performances satisfaisantes avec des scores élevés de précision, de rappel et de score F1 pour la plupart des classes de la catégorie "Geometry". Cependant, certaines classes montrent des performances légèrement inférieures.

5.8 Architecture VGG16

Le VGG16 (Visual Geometry Group), est un réseau neuronal convolutif profond. Sa structure, composée de treize couches de convolution et trois couches entièrement connectées, a démontré une excellente qualité de résultats en classification. Ces couches convolutives sont organisées en cinq blocs, avec une taille de filtre uniforme de 3x3. Les deux premiers blocs ont deux couches convolutives chacun, tandis que les trois derniers blocs ont trois à quatre couches. Après chaque bloc, une couche de Maxpooling réduit la dimension spatiale des cartes de caractéristiques.

Ce modèle est conçu pour fonctionner avec des images d'entrée normalisées, présentées sous forme de mini-lots d'images RVB à 3 canaux. Les images doivent avoir une dimension (3 × H × W), avec les valeurs de H et W fixées à 224. Afin d'améliorer le modèle VGG16, les couches de classification finales ont été remplacées par des couches entièrement connectées comme illustré figure 5.8. Les fonctions d'activation ReLU et softmax ont été utilisées pour ajouter des non-linéarités au modèle et pour obtenir des probabilités de classe pour les prédictions. Pour préserver les connaissances acquises lors de la formation d'ImageNet, les

Les poids de couches initiales du modèle, qui correspondent aux couches convolutives pré-formées, ont été maintenues figées. La couche modifiée a été spécifiquement affinée avec le jeu de données B-Scans.

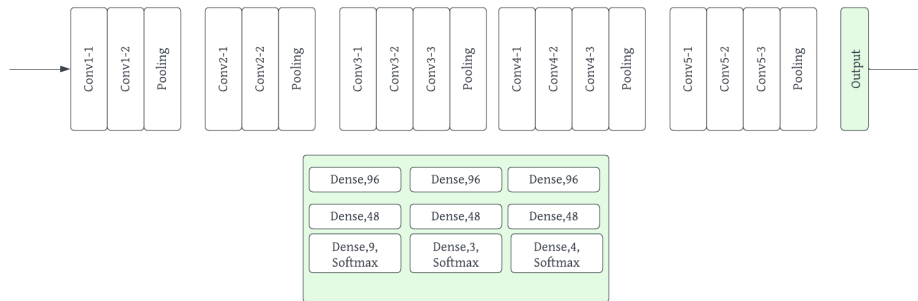
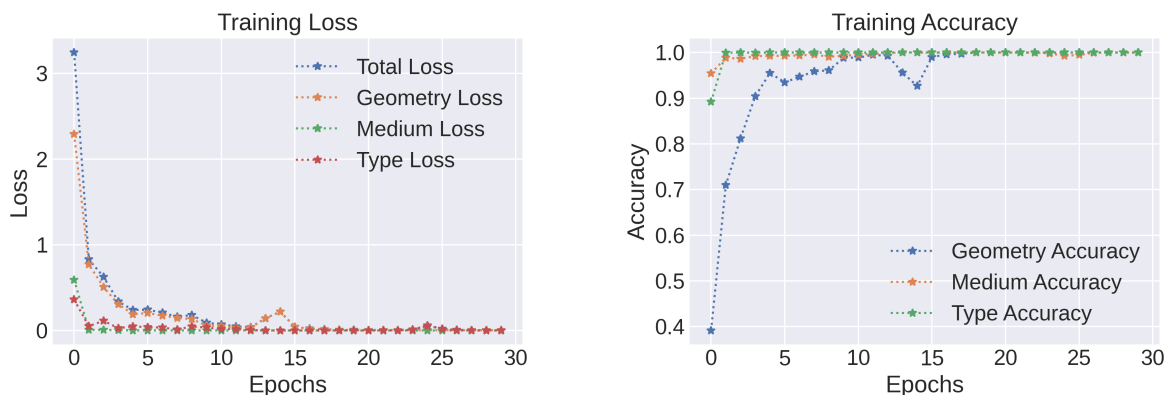


Figure 5.17: Stratégie globale de transfert par apprentissage.

Au cours de 30 époques, le modèle a été formé avec une taille de lot de 64. Afin de minimiser la perte d'entropie croisée catégorielle, l'optimiseur Adam a été utilisé pour ajuster les poids du modèle. Diverses méthodes ont été employées pour améliorer la vitesse de calcul et les performances globales du modèle. Les techniques employées dans cette approche consistent en un arrêt précoce pour éviter le surajustement (overfitting). Une augmentation des données permet d'enrichir l'ensemble d'apprentissage avec plus de diversité.



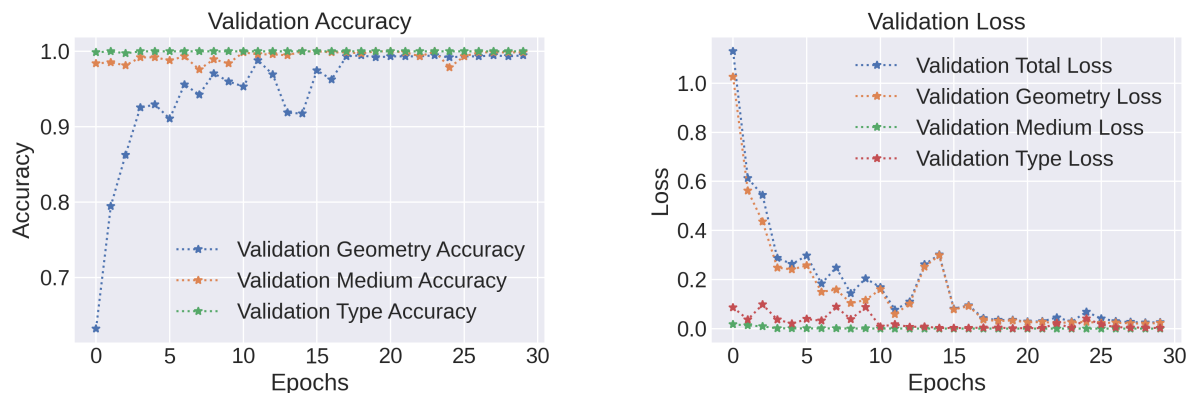


Figure 5.18: Les résultats de VGG-16 représentant l'évolution de l'entraînement et de la validation (a) de la perte et (b) de l'exactitude au fil des époques.

En ce qui concerne l'évolution de la perte sur les ensembles d'entraînement et de validation, le modèle VGG16 a démontré une tendance constante à la diminution de la fonction de perte, aussi bien pour l'ensemble d'entraînement que pour l'ensemble de validation, au fil des époques. Au début de l'entraînement, la perte sur les ensembles d'entraînement et de validation était relativement élevée, mais elle a très rapidement diminué au fil des époques jusqu'à converger vers une valeur proche de zéro. Cependant, entre les époques 10 et 15, on observe une certaine instabilité dans la perte, avec des fluctuations significatives. Cette instabilité est générée par la mise à jour des poids. La convergence indique que le modèle VGG16 a réussi à ajuster ses paramètres de manière à minimiser la différence entre les prédictions et les valeurs réelles.

Concernant l'historique de précision sur l'ensemble d'entraînement et de validation, celui-ci subit une augmentation constante au fil des époques sur l'ensemble d'entraînement et de validation. Initialement, la précision était relativement faible, mais elle a rapidement augmenté au fur et à mesure que le modèle apprenait et s'ajustait aux données. En outre, entre les 10e et 15e époques, l'apprentissage a été retardé et limité en raison de l'actualisation du taux d'apprentissage à des valeurs optimales plus faibles pour calculer le meilleur résultat. Au fil des époques, la précision a continué à augmenter, atteignant éventuellement une valeur élevée proche de 1.00 (ou 100%) pour les deux ensembles. Cette augmentation constante de la précision indique que le modèle VGG16 a été capable d'apprendre efficacement les caractéristiques des images des B-Scan.

Dans l'ensemble, les résultats du modèle VGG16 sur les courbes de pertes et de précision démontrent un bon ajustement du modèle aux données d'entraînement

et de validation. La perte diminue régulièrement et converge vers une valeur faible. De même, la précision augmente continuellement pour atteindre une valeur élevée. Ces observations indiquent que le modèle VGG16 est capable de capturer les motifs importants des données et de faire des prédictions précises pour l'ensemble des trois classes d'étiquetage.

Table 5.8: Rapport de classification

Rapport de classification	Précision	Recall	F1-Score
diamètre 16 mm	1.00	1.00	1.00
diamètre 24 mm	1.00	1.00	1.00
diamètre 32 mm	1.00	1.00	1.00
diamètre 40 mm	1.00	0.98	0.99
diamètre 48 mm	0.99	1.00	0.99
diamètre 64 mm	1.00	1.00	1.00
diamètre 72 mm	1.00	1.00	1.00
diamètre 80 mm	1.00	1.00	1.00
diamètre 100 mm	1.00	1.00	1.00
Macro average	1.00	1.00	1.00
Weighted average	1.00	1.00	1.00

En résumé, les résultats des rapports de classification indiquent que le modèle a atteint une performance parfaite pour toutes les classes dans les ensembles de test, avec une précision, un rappel et un score F1 de 1.00. Cela suggère que le modèle a appris efficacement les caractéristiques discriminantes de chaque classe et est capable de faire des prédictions précises pour de nouvelles instances. L'exactitude globale de 1.00 confirme la capacité du modèle à classer correctement toutes les instances des ensembles de test pour chaque étiquette.

5.9 CNN personnalisé

On rappelle que dans ce travail, un réseau de neurones convolutifs profond est développé pour la classification multi-étiquette afin de classer 9 diamètres différents, 3 types de tuyaux différents et 4 milieux de propagation différents. Le modèle CNN personnalisé s'inspire du réseau Inception développé par Google [47]. La principale différence entre le modèle Inception et les CNN ordinaires réside dans les blocs Inception. Ces blocs consistent à convoluer la même entrée avec des filtres de taille différente et à concaténer leurs résultats. L'objectif de ce modèle est d'introduire le concept de Multi-pathing, qui permet de capturer et

d'extraire simultanément différentes caractéristiques. Le bloc de construction est illustré dans la figure , ce bloc contient différentes couches de convolution avec des tailles de filtre différentes (1x1), (5x5), (7x7). La diversité des tailles de noyau permet d'augmenter la capacité du réseau à extraire les caractéristiques les plus complexes. 32 filtres de taille 3×3 sont utilisés dans la couche d'entrée, suivis d'une couche de MaxPooling, puis d'une couche de normalisation par lots pour augmenter la vitesse et la stabilité de l'apprentissage. Une couche de convolution avec 64 filtres de taille 3×3 est ensuite définie. Cette couche est précédée d'une couche de MaxPooling, suivie de deux blocs de construction Inception représentés ci-dessus. Chaque couche de convolution contient respectivement 128 et 16 filtres de taille (1x1), (5x5), (7x7). A la fin de chaque bloc Inception, toutes les cartes de caractéristiques sont concaténées. La fonction d'activation utilisée dans toutes les couches de convolution est la fonction ReLU. Ensuite, toutes les cartes de caractéristiques sont transformées en un vecteur unidimensionnel à l'aide d'une couche aplatie. Un ensemble de 3 couches denses a été défini, contenant respectivement 256, 512 et 256 neurones. Une fonction d'activation ReLU est sélectionnée pour chaque couche. Enfin, 3 branches ont été définies pour chaque sortie, chaque branche comprenant en amont un ensemble de trois couches denses composées respectivement de 96, 48 et 48 neurones. Chaque couche de sortie comporte 9 neurones permettant de reconnaître la taille du diamètre, 4 neurones pour spécifier le milieu de propagation et 3 neurones pour distinguer le type de tuyau. Les couches de sortie sont suivies d'une couche softmax principalement utilisée pour les problèmes de classification multi-étiquette afin de prédire la probabilité de chaque étiquette.

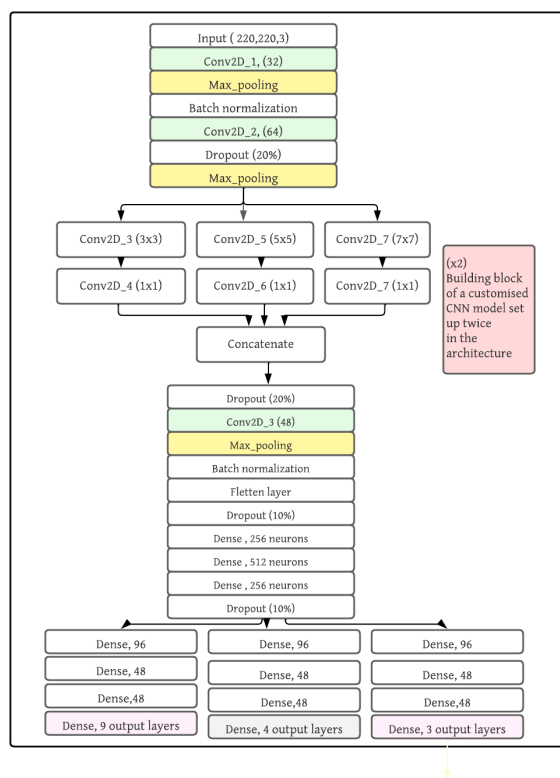


Figure 5.19: Architecture CNN de référence

Dans le cadre de notre étude, nous n'intégrons pas les résultats détaillés à ce stade, mais notre focus principal réside dans l'exploration de l'optimisation bayésienne pour l'étude des hyperparamètres du modèle en question. Nous prévoyons de détailler et de discuter les subtilités de l'optimisation des hyperparamètres de ce modèle dans les sections suivantes, offrant ainsi un aperçu précis de la manière dont chaque hyperparamètre influence la performance du modèle.

5.10 Influence des hyperparamètres

L'objectif de cette section est d'étudier l'influence de chaque hyperparamètre. Pour ce faire, nous utiliserons l'optimisation bayésienne, une technique basée sur la théorie des probabilités, afin de maximiser la précision de notre modèle CNN personnalisé.

Dans cette étude, nous détaillerons notre méthodologie, y compris la structure du modèle CNN que nous utilisons comme base, ainsi que la définition de l'espace de recherche pour les hyperparamètres. Nous présenterons ensuite les résultats

de l'optimisation, en mettant en évidence les configurations les plus performantes. Rappelons que notre modèle CNN a une architecture complexe avec plusieurs couches de convolutions, des branchements, et de couches de sortie multiples, le modèle est conçu pour traiter des images de taille 220x220x3.

5.10.1 Méthodologie

Dans cette section, nous présentons de manière plus détaillée la méthodologie suivie pour la sélection des hyperparamètres. Pour ce faire, nous commençons tout d'abord par définir l'espace des hyperparamètres.

Parmi les principaux hyperparamètres, nous avons sélectionné la liste suivante :

- Le nombre de filtres dans les couches de convolution.
- La taille du noyau de convolution.
- Le taux d'abandon (dropout rate).
- Le nombre d'époques d'entraînement.
- La taille du lot (batch size).

Afin de déterminer les meilleures combinaisons d'hyperparamètres, nous avons opté pour l'optimisation bayésienne, technique connue pour ses supériorités face à des méthodes plus conventionnelles comme la recherche exhaustive (GridSearch). En contraste avec GridSearch, qui teste toutes les combinaisons possibles dans l'espace de variation défini des hyperparamètres, l'optimisation bayésienne sélectionne stratégiquement les régions de cet espace susceptibles de présenter les meilleures performances, optimisant ainsi l'efficacité et l'économie des ressources de calcul.

Nous avons utilisé la bibliothèque Hyperopt, basée sur l'algorithme Tree-structured Parzen Estimator (TPE), pour mettre en œuvre cette technique d'optimisation, due à sa réputation dans l'efficacité d'exploration des espaces d'hyperparamètres.

Le protocole commence par la conception d'une fonction objectif qui évalue diverses combinaisons d'hyperparamètres et renvoie une métrique de perte associée. Notre objectif est de minimiser la perte de validation de notre modèle de réseau de neurones convolutif (CNN), car cette métrique sert de reflet de la performance du modèle sur un ensemble de données distinct de celui d'entraînement, cruciale pour éviter l'overfitting.

L'optimisation bayésienne, de nature itérative, vise, à chaque itération, à examiner l'espace des hyperparamètres afin d'identifier les configurations optimales. L'algorithme, dans notre cas, procède à travers 20 essais, exploitant les données

recueillies à chaque étape pour perfectionner la recherche. Tous les résultats, y compris les métriques de perte et les performances du modèle, sont conservés pour des analyses futures.

Finalement, une fois la phase d'optimisation achevée, la configuration d'hyperparamètres induisant la moindre perte de validation est sélectionnée. Bien que l'architecture du réseau demeure constante, le choix judicieux des hyperparamètres s'avère vital pour la mise au point d'un modèle qui répond parfaitement à nos exigences spécifiques.

L'optimisation des poids d'un réseau de neurones implique l'utilisation de l'algorithme de rétropropagation avec la descente de gradient stochastique. Parmi les hyperparamètres essentiels de cette méthode, nous trouvons le taux d'apprentissage α_0 et le coefficient γ . Le coefficient γ est crucial, car il sert à moduler la manière dont le taux d'apprentissage α décroît au fil des époques pendant l'entraînement. En d'autres termes, γ détermine à quel rythme l'adaptation des poids du réseau de neurones se fait, permettant ainsi un contrôle fin de la convergence du modèle pendant l'entraînement. Entraîner le réseau de neurones avec les valeurs optimales de α_0 et γ permet d'obtenir un modèle de classification performant. L'approche consiste à commencer avec un taux d'apprentissage élevé α_0 et à le réduire progressivement par un facteur multiplicatif constant γ pendant l'entraînement, jusqu'à la stabilisation de l'erreur de validation ou l'arrêt de la décroissance de l'erreur d'apprentissage.

$$\alpha_{i+1} = \alpha_i \cdot (1 + \gamma)^{-1} \quad (5.11)$$

Cette équation montre comment le taux d'apprentissage α est ajusté à chaque époque en fonction de la valeur de γ , assurant ainsi une décroissance contrôlée du taux d'apprentissage et, par conséquent, une convergence plus stable et efficace du modèle.

Afin de sélectionner les hyperparamètres optimaux pour le taux d'apprentissage α_0 et le coefficient γ , nous allons explorer différentes combinaisons de valeurs pour ce couple, notamment :

- Taux d'apprentissage α_0 avec les valeurs suivantes : 10^{-3} , 5×10^{-3} , 10^{-2} .
- Coefficient γ avec les valeurs suivantes : 10^{-4} , 10^{-3} , 10^{-2} , 10^{-1} . Pour éviter une convergence lente, il est fortement recommandé de commencer avec une petite valeur (0,01) et de l'ajuster en conséquence.

Meilleure combinaisons pour la catégorie diamètre:

Les deux tableaux 5.11, 5.10 présentent les meilleures combinaisons d'hyperparamètres

Architecture	Taux d'apprentissage α_0	Coefficient γ
Cnn	10^{-3}	10^{-3}

Table 5.9: Meilleurs hyper-paramètres sélectionnés

Métrique	Valeur
Filter	32.000000
Kernel_Size	5.000000
Dropout_Rate	0.000000
Learning_Rate	0.0010000
Epoch	15.000000
Batch_Size	32.000000

Table 5.10: Combinaison des hyperparamètres optimisés: Classes milieu de propagation et type de tuyau.

pour deux catégories distinctes: le "Milieu de propagation et type de tuyau" et le "diamètre".

En comparant les deux tableaux:

- Filtre : Pour les deux catégories, le nombre optimal de filtres est le même, soit 32.
- Taille du noyau / Kernel Size: Pour la catégorie "Milieu de propagation et type de tuyau", une taille de noyau de 5 est optimale, tandis que pour la catégorie "diamètre", une taille de noyau de 7 est préférée.
- Taux d'abandon / Dropout Rate: Le taux d'abandon est de 0 pour les deux catégories, indiquant qu'il n'y a pas de désactivation d'unités pendant l'entraînement.
- Taux d'apprentissage / Learning Rate: Le taux d'apprentissage optimal est identique pour les deux catégories, soit 0.001.
- Taille du lot / Batch Size: Pour la catégorie "Milieu de propagation et type de tuyau", une taille de lot de 32 est optimale. En revanche, pour la catégorie "diamètre", une taille de lot beaucoup plus grande, soit 128, est préférée.
- Époques / Epoch: Le nombre d'époques est le même pour les deux, fixé à 15 époques.
- Taille du lot / Batch Size: Pour la catégorie "Milieu de propagation et type de tuyau", une taille de lot de 32 est optimale. En revanche, pour la catégorie "diamètre", une taille de lot beaucoup plus grande, soit 128, est optimale.

En résumé, les hyperparamètres sont généralement très similaires entre les deux catégories, à l'exception de la taille du noyau et de la taille du lot.

Métrique	Valeur
filter_	32
kernel_size	7
dropout_rate	0.10
Learning_Rate	0.001
epoch	15
batch_size	128

Table 5.11: Combinaison des hyperparamètres optimisés: Classe diamètres.

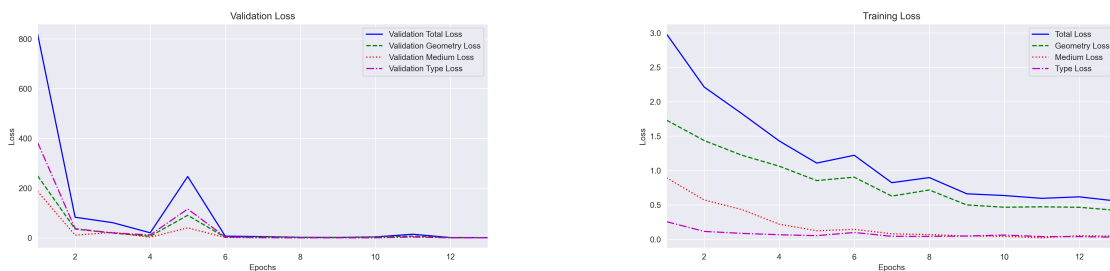


Figure 5.20: Résultats historiques pour les pertes de validation et d'entraînement.

Les figures 5.20 montre la progression de l'apprentissage de notre modèle CNN sur 13 époques (utilisation d'early stopping), indiquant une perte totale et trois pertes spécifiques : Géométrie, Medium et Type. Globalement, toutes les pertes tendent à diminuer au fil des époques, signe d'une amélioration de la performance du modèle. Lors de la comparaison des résultats d'entraînement et de validation, le modèle a montré une forte suradaptation initiale avec une perte de validation bien supérieure. Au fil des époques, les deux pertes convergent, indiquant une amélioration de la généralisation.

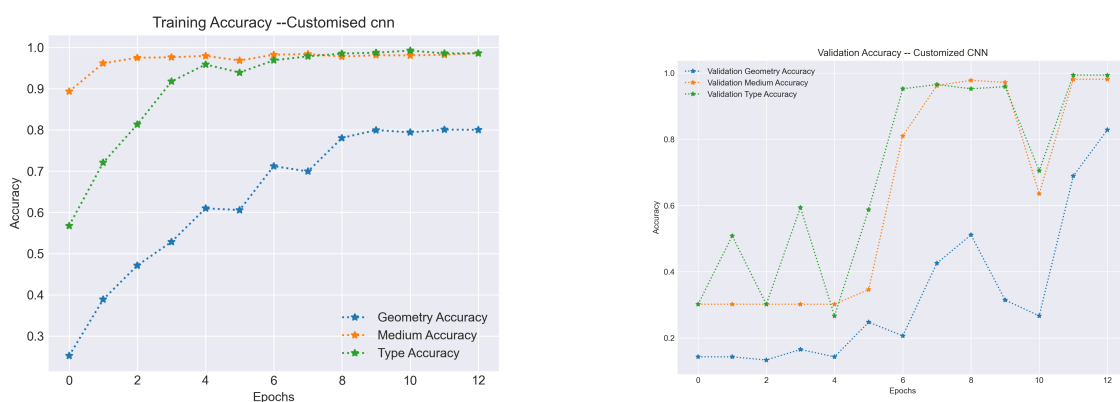


Figure 5.21: Résultats historiques pour la précision de validation et d'entraînement.

En comparant les deux résultats historiques pour la précision, de validation et d'entraînement, on observe que

- En ce qui concerne la précision de géométrie, les valeurs de précision de validation initiales sont faibles, démarrant à environ 14.29% à la première époque, illustrant ainsi une certaine instabilité ou dans les performances du modèle durant les premières époques. Cependant, une amélioration significative est observée au fil des époques, la précision atteignant 82,9% à la 13ème époque.
- Pour la précision du milieu: la précision de validation montre une amélioration remarquable au fil du temps, passant de 30,2% à 98,1%. La précision d'entraînement commence à un niveau élevé de 89,3% et se stabilise autour de 98,6%.
- Pour la précision du type de tuyau : la précision de validation présente des fluctuations mais montre une amélioration globale, passant de 30,2% à 99,4%. Les précisions d'entraînement augmentent constamment, commençant à 56,8% et culminant à 98,7%.

Globalement, le modèle semble avoir de meilleures performances en termes d'entraînement que de validation. Cependant, vers la fin, les performances de validation se rapprochent des performances d'entraînement, indiquant une amélioration de la généralisation.

Afin d'effectuer une évaluation approfondie de notre modèle , nous avons calculé un rapport de classification qui détaille les performances pour chaque classe:

Table 5.12: Rapport de classification sur les données de test- classe diamètre

Label	Précision	Rappel	Score F1
diam100	0.95	1.00	0.98
diam16	0.92	1.00	0.96
diam24	0.98	0.69	0.81
diam32	0.83	0.79	0.76
diam40	0.84	0.87	0.80
diam48	0.84	0.75	0.79
diam64	1.00	0.90	0.94
diam72	0.85	0.95	0.84
diam80	0.88	0.77	0.82
accuracy			0.86
macro avg	0.89	0.86	0.86
weighted avg	0.90	0.86	0.86

Ce tableau 5.12 présente un rapport de classification pour différents diamètres. En observant les valeurs de précision, rappel et score F1, il apparaît que les classes "diam100" et "diam16" ont des performances exceptionnellement élevées avec une précision et un rappel à ou près de 1.00. Globalement, la précision moyenne, le rappel moyen et le score F1 pour l'ensemble des classes se situent autour de 0.86 à 0.90, indiquant une performance solide du modèle sur l'ensemble des données de test.

Table 5.13: Rapport de classification classe milieu de propagation

Label	Précision	Rappel	Score F1
medium_labels_argile	1.00	1.00	1.00
medium_labels_beton	1.00	0.98	0.99
medium_labels_fractal	0.98	0.99	0.99
medium_labels_sable	0.99	1.00	1.00
accuracy			0.99
macro avg	0.99	0.99	0.99
weighted avg	0.99	0.99	0.99

Le tableau 5.13 montre une performance exceptionnelle des modèles sur l'ensemble de test avec une précision globale de 99%. En particulier, les labels "argile" et "sable" ont atteint une perfection avec un score F1 de 1,00, tandis que 'béton' et 'fractal' se rapprochent également de cette performance optimale."

Table 5.14: Rapport de classification pour les labels de type de tuyau

Label	Précision	Rappel	Score F1
typeOfPipe_labels_eau	1.00	0.95	0.97
typeOfPipe_labels_metal	1.00	1.00	1.00
typeOfPipe_labels_vide	0.96	1.00	0.98
accuracy			0.98
macro avg	0.99	0.98	0.98
weighted avg	0.99	0.98	0.98

Le modèle a parfaitement identifié les tuyaux de type "métal" avec une précision et un score F1 de 1,00. Les tuyaux de type "eau" ont été légèrement moins bien classifiés, avec une précision parfaite, mais un rappel de 0,95. Les tuyaux de type "vide" ont également obtenu d'excellents résultats, avec un rappel parfait et une légère diminution de la précision à 0,96. Globalement, l'exactitude du modèle sur l'ensemble de test est de 98%, ce qui témoigne de sa robustesse et de sa fiabilité.

5.11 Synthèse

Dans ce chapitre, nous nous sommes attachés à explorer l'application des réseaux de neurones profonds à l'analyse de radars à sondage de sol, en se focalisant particulièrement sur la classification d'objets enfouis, à partir d'images synthétiques obtenues par la méthode FDTD.

Le développement de notre base de données s'est opéré sous des configurations spécifiques, où l'émetteur est fixe tandis que le récepteur se déplace selon une direction déterminée sur l'interface air/sol, permettant de réaliser un déplacement sur une distance de 0.78 m en générant les A-scans nécessaires à la création du B-scan. Une attention particulière a été portée à la division de la base de données en ensembles distincts d'apprentissage et de test.

Nous avons approfondi le prétraitement des données, une phase incontournable dans la réalisation de modèles d'apprentissage profond. Ce processus englobe l'extraction de caractéristiques, la segmentation de la base de données, l'entraînement des modèles de CNN, et la validation des résultats acquis. La détection de contours, utilisant des méthodes telles que l'opérateur de Canny, a joué un rôle important dans le traitement d'images, permettant la délimitation précise des objets dans l'ensemble des images radagrammes.

Trois architectures de réseaux de neurones profonds ont été déployées et leurs performances ont été comparées à travers divers métriques, pertinentes à la problématique de classification Multi-label.

L'analyse comparative des trois architectures pour la classification des différents diamètres est la suivante:

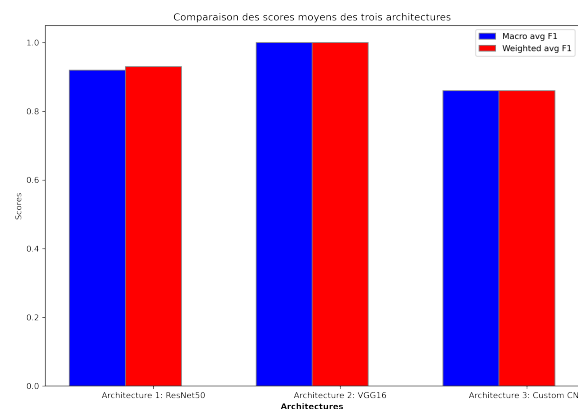


Figure 5.22: Comparaison des scores moyens des trois architectures

- **Première architecture ResNet-50:** Elle présente des performances variables selon les classes, avec des scores F1 oscillant entre 0,76 pour le diamètre 72mm et 0,98 pour le diamètre 100mm et 32mm. Les scores F1 moyens pondérés et macro sont respectivement de 0,93 et 0,92, ce qui indique une performance globalement bonne. Toutefois, certains diamètres (comme 72mm et 80mm) pourraient nécessiter une attention supplémentaire pour amélioration.
- **Deuxième architecture VGG-16:** Cette architecture se démarque par sa performance exceptionnelle. Presque toutes les classes ont un score F1 parfait de 1,00, à l'exception du diamètre 40mm qui est légèrement en dessous à 0,99. Les scores F1 moyens pondérés et macro sont tous les deux à 1,00, témoignant de l'excellente performance globale de cette architecture.
- **Troisième architecture Custom CNN:** Elle présente des scores F1 allant

de 0,76 pour le diamètre 32mm à 0,98 pour le diamètre 100mm. Le score d'exactitude global est de 0,86, ce qui est légèrement inférieur à celui de la première architecture. Les scores F1 moyens pondérés et macro sont respectivement de 0,86 et 0,89, indiquant une performance globalement bonne mais offrant une marge d'amélioration.

La deuxième architecture est clairement supérieure en termes de performance, atteignant presque la perfection sur toutes les classes. La première et la troisième architecture sont compétitives, mais la première semble avoir une légère avance en termes de scores F1 moyens.

Conclusion Générale et perspectives

Après avoir décrit les notions utilisées pour construire des réseaux de neurones profonds, cette thèse explore le développement de nouvelles méthodes de diagnostic pour identifier et situer les défauts dans les systèmes de réseaux électriques complexes. Cette problématique constitue un enjeu majeur pour les fabricants. L'importance et la complexité des défauts non-francs ont été mises en avant et détaillées, tout en exposant les insuffisances des techniques existantes pour une topologie complexe. Par conséquent, le deep learning a été identifié comme un outil efficace pour le diagnostic de ces défauts.

Dans la première partie, nous nous sommes concentrés sur le déploiement du deep learning pour améliorer la détection et la localisation des défauts dans les réseaux de câbles. Les architectures de réseaux de neurones, telles que le Multi-Layer Perceptron (MLP) et le Convolutional Neural Network (CNN), ont été explorées et optimisées, fournissant un cadre solide et robuste pour améliorer la précision et l'efficacité dans la gestion des défauts non-francs au sein des systèmes câblés.

Un second focus de notre recherche a été de mettre en application les réseaux de neurones profonds pour analyser les données du radar à pénétration de sol (GPR), avec une attention particulière portée sur la classification des objets enfouis. Nous avons utilisé des images synthétiques, obtenues par la méthode FDTD.

Le développement de notre base de données a été effectué avec une configuration précise où l'émetteur reste statique et le récepteur se déplace selon une direction prédéterminée à l'interface air/sol. Cela nous a permis d'effectuer l'exploration sur une distance de 0,78 m tout en générant les A-scans requis pour la création du B-scan. L'étape de prétraitement des données a été approfondie, puisqu'elle est indispensable à la réalisation des modèles de deep learning. Cela a inclus l'extraction de caractéristiques, la segmentation de la base de données, l'entraînement des modèles de CNN (réseaux de neurones convolutionnels), et la validation des résultats obtenus. Le processus de détection des contours, au moyen de méthodes telles que l'opérateur de Canny, a été d'une importance cruciale pour le traitement des images, en permettant une délimitation nette des objets dans l'ensemble des images radargrammes.

Nous avons déployé trois architectures de réseaux de neurones profonds et comparé leurs performances à l'aide de diverses métriques, pertinentes pour le problème de classification multi-label. Cette étape de comparaison nous a permis

d'évaluer l'efficacité de chaque architecture dans le contexte de nos données et objectifs spécifiques, nous guidant ainsi vers le modèle le plus performant pour notre application spécifique. Ces méthodes de deep learning se sont avérées non seulement essentielles pour gérer les données massives et complexes du GPR mais aussi pour améliorer considérablement la fiabilité et la précision de nos méthodes d'analyse.

En somme, l'application de réseaux de neurones profonds à l'analyse de données GPR s'est avérée être une approche efficace, améliorant significativement notre capacité à détecter et à caractériser des objets enfouis, et offrant une méthode robuste et précise d'analyse des données radar.

Perspectives

Les travaux menés au cours de cette thèse ouvrent de nombreuses pistes pour la recherche future, en particulier dans le développement et la validation ultérieure des modèles de deep learning dans le contexte de diagnostic des câbles et de l'analyse des données du radar à pénétration de sol (GPR).

Génération et utilisation de données bruitées: La robustesse des modèles développés pourrait être davantage mise à l'épreuve en les testant avec des données bruitées, ce qui simulerait des conditions réelles plus fidèlement. Cependant, le processus de génération de données via la méthode FDTD est long. Ainsi, l'exploration de méthodes alternatives ou optimisées pour la génération de données bruitées pourrait s'avérer précieuse.

Exploration des réseaux LSTM pour l'analyse des signaux GPR: Les réseaux Long Short-Term Memory (LSTM), étant particulièrement performants pour le traitement des données séquentielles donc pour les signaux temporels, pourraient être utilisés pour améliorer l'analyse des signaux A-scan dans les données GPR. Cela pourrait permettre de capturer efficacement les dépendances temporelles et les caractéristiques sous-jacentes des signaux, améliorant ainsi la précision de la détection des objets enfouis.

Application des réseaux LSTM pour l'analyse des signaux de réflectométrie: Concernant la partie réflectométrie, l'utilisation de LSTM pour analyser l'intégralité des signaux de réflectométrie offre une voie prometteuse. Les LSTM pourraient potentiellement capturer des motifs plus complexes dans les données temporelles, ce qui pourrait se traduire par une meilleure identification et localisation des défauts.

Validation des modèles avec des données expérimentales: Bien que les modèles aient été développés et validés à l'aide de données synthétiques, une étape supplémentaire critique serait de les tester et de les valider à l'aide de données expérimentales réelles. Cela permettrait d'assurer que les modèles sont capables de généraliser leurs apprentissages à de nouvelles données et sont applicables dans des scénarios du monde réel.

Références

- [1] M. Minsky, *Society of mind*. Simon and Schuster, 1988.
- [2] A. M. Turing, *Computing machinery and intelligence*. Springer, 2009.
- [3] A. M. Turing, *The essential turing*. Oxford University Press, 2004.
- [4] R. S. Michalski, J. G. Carbonell, and T. M. Mitchell, *Machine learning: An artificial intelligence approach*. Springer Science & Business Media, 2013.
- [5] A. L. Samuel, "Machine learning," *The Technology Review*, vol. 62, no. 1, pp. 42–45, 1959.
- [6] C.-A. Azencott, *Introduction au Machine Learning-2e éd*. Dunod, 2022.
- [7] P. Domingos, "A few useful things to know about machine learning," *Communications of the ACM*, vol. 55, no. 10, pp. 78–87, 2012.
- [8] S. Rong and Z. Bao-Wen, "The research of regression model in machine learning field," in *MATEC Web of Conferences*, EDP Sciences, vol. 176, 2018, p. 01 033.
- [9] V. Gulshan, L. Peng, M. Coram, *et al.*, "Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs," *jama*, vol. 316, no. 22, pp. 2402–2410, 2016.
- [10] M. S. Sorower, "A literature survey on algorithms for multi-label learning," *Oregon State University, Corvallis*, vol. 18, no. 1, p. 25, 2010.
- [11] J. Brownlee, *Machine learning mastery with Python: understand your data, create accurate models, and work projects end-to-end*. Machine Learning Mastery, 2016.
- [12] Y. Bengio, Y. Lecun, and G. Hinton, "Deep learning for ai," *Commun. ACM*, vol. 64, no. 7, pp. 58–65, Jun. 2021. [Online]. Available: <https://doi.org/10.1145/3448250>.
- [13] J. Brownlee, "Supervised and unsupervised machine learning algorithms," *Machine Learning Mastery*, vol. 16, no. 03, 2016.

- [14] M. Sokolova and G. Lapalme, “A systematic analysis of performance measures for classification tasks,” *Information processing & management*, vol. 45, no. 4, pp. 427–437, 2009.
- [15] D. M. Powers, “Evaluation: from precision, recall and f-measure to roc, informedness, markedness and correlation,” *arXiv preprint arXiv:2010.16061*, 2020.
- [16] A. Botchkarev, “Evaluating performance of regression machine learning models using multiple error metrics in azure machine learning studio,” *Available at SSRN 3177507*, 2018.
- [17] C. Touzet, *les réseaux de neurones artificiels, introduction au connexionnisme*. Ec2, 1992.
- [18] W. S. McCulloch and W. Pitts, “A logical calculus of the ideas immanent in nervous activity,” *Bulletin of mathematical biology*, vol. 52, pp. 99–115, 1990.
- [19] G. Shaw, “Donald hebb: the organization of behavior,” in *Brain Theory: Proceedings of the First Trieste Meeting on Brain Theory, October 1–4, 1984*, Springer, 1986, pp. 231–233.
- [20] F. Rosenblatt *et al.*, *Principles of neurodynamics: Perceptrons and the theory of brain mechanisms*. Spartan books Washington, DC, 1962, vol. 55.
- [21] M. Minsky and S. Papert, “An introduction to computational geometry,” *Cambridge tiass., HIT*, vol. 479, no. 480, p. 104, 1969.
- [22] P. J. Werbos, “Backpropagation through time: what it does and how to do it,” *Proceedings of the IEEE*, vol. 78, no. 10, pp. 1550–1560, 1990.
- [23] Y. LeCun *et al.*, “Lenet-5, convolutional neural networks,” *URL: <http://yann.lecun.com/exdb/lenet>*, vol. 20, no. 5, p. 14, 2015.
- [24] S. Hochreiter and J. Schmidhuber, “Lstm can solve hard long time lag problems,” *Advances in neural information processing systems*, vol. 9, 1996.
- [25] M. Riedmiller and A. Lernen, “Multi layer perceptron,” *Machine Learning Lab Special Lecture, University of Freiburg*, pp. 7–24, 2014.
- [26] Y. Le Cun, “Learning process in an asymmetric threshold network,” in *Disordered systems and biological organization*, Springer, 1986, pp. 233–240.
- [27] J. Schmidhuber, “Jürgen schmidhuber,” *Deep learning in neural networks: An overview*, *Neural Networks*, vol. 61, pp. 85–117, 2015.

- [28] S. Sharma, S. Sharma, and A. Athaiya, “Activation functions in neural networks,” *Towards Data Sci*, vol. 6, no. 12, pp. 310–316, 2017.
- [29] F. Chollet, *Deep learning with Python*. Simon and Schuster, 2021.
- [30] C. Khosla and B. S. Saini, “Enhancing performance of deep learning models with different data augmentation techniques: a survey,” in *2020 International Conference on Intelligent Engineering and Management (ICIEM)*, IEEE, 2020, pp. 79–85.
- [31] I. Goodfellow, Y. Bengio, and A. Courville, “Regularization for deep learning,” *Deep learning*, pp. 216–261, 2016.
- [32] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, “Dropout: a simple way to prevent neural networks from overfitting,” *The journal of machine learning research*, vol. 15, no. 1, pp. 1929–1958, 2014.
- [33] Y. Bai, E. Yang, B. Han, *et al.*, “Understanding and improving early stopping for learning with noisy labels,” *Advances in Neural Information Processing Systems*, vol. 34, pp. 24 392–24 403, 2021.
- [34] R. Kohavi *et al.*, “A study of cross-validation and bootstrap for accuracy estimation and model selection,” in *Ijcai*, Montreal, Canada, vol. 14, 1995, pp. 1137–1145.
- [35] S. Arlot and A. Celisse, “A survey of cross-validation procedures for model selection,” 2010.
- [36] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, “Gradient-based learning applied to document recognition,” *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.
- [37] A. A. A. Shervine, *Pense-bête de réseaux de neurones convolutionnels*.
- [38] S.-i. Amari, “Backpropagation and stochastic gradient descent method,” *Neurocomputing*, vol. 5, no. 4-5, pp. 185–196, 1993.
- [39] I. Sutskever, J. Martens, G. Dahl, and G. Hinton, “On the importance of initialization and momentum in deep learning,” in *International conference on machine learning*, PMLR, 2013, pp. 1139–1147.
- [40] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, “Learning representations by back-propagating errors,” *nature*, vol. 323, no. 6088, pp. 533–536, 1986.
- [41] S. Nedjaoume, “Entrai nement d’un réseau de neurones mlp par la méthode bfgs,” 2016.
- [42] M. L. Lenard, “Practical convergence conditions for the davidon-fletcher-powell method,” *Mathematical Programming*, vol. 9, pp. 69–86, 1975.

- [43] J. Duchi, E. Hazan, and Y. Singer, “Adaptive subgradient methods for online learning and stochastic optimization.,” *Journal of machine learning research*, vol. 12, no. 7, 2011.
- [44] M. D. Zeiler, “Adadelata: an adaptive learning rate method,” *arXiv preprint arXiv:1212.5701*, 2012.
- [45] D. P. Kingma and J. Ba, “Adam: a method for stochastic optimization,” *arXiv preprint arXiv:1412.6980*, 2014.
- [46] S. Tammina, “Transfer learning using vgg-16 with deep convolutional neural network for classifying images,” *International Journal of Scientific and Research Publications (IJSRP)*, vol. 9, no. 10, pp. 143–150, 2019.
- [47] A. Khan, A. Sohail, U. Zahoor, and A. S. Qureshi, “A survey of the recent architectures of deep convolutional neural networks,” *Artificial intelligence review*, vol. 53, pp. 5455–5516, 2020.
- [48] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “Imagenet classification with deep convolutional neural networks,” *Communications of the ACM*, vol. 60, no. 6, pp. 84–90, 2017.
- [49] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.
- [50] F. Auzanneau, “Wire troubleshooting and diagnosis: review and perspectives,” *Progress In Electromagnetics Research B*, vol. 49, pp. 253–279, 2013.
- [51] P. Tsai, C. Lo, Y. C. Chung, and C. Furse, “Mixed-signal reflectometer for location of faults on aging wiring,” *IEEE Sensors Journal*, vol. 5, no. 6, pp. 1479–1482, 2005.
- [52] C. Furse, “Finding fault: locating hidden hazards on aircraft wiring,” *College of Engineering, University of Utah*, 2004.
- [53] P. B. Ladkin and W. Schepper, “Emi, twa 800 and swissair 111,” *The resks digist*, vol. 21, 2000.
- [54] C. Wilwert, N. Navet, F. Simonot-Lion, *et al.*, *Design of automotive x-by-wire systems*, 2005.
- [55] F. Pullaczek, “Théorie du câble coaxial. i,” *Journal de Physique et le Radium*, vol. 8, no. 7, pp. 215–224, 1947.
- [56] A. Knobloch, H. Garbe, and J. Karst, “Shielded or unshielded twisted-pair for high speed data transmission?” In *1998 IEEE EMC Symposium*.

- International Symposium on Electromagnetic Compatibility. Symposium Record (Cat. No. 98CH36253)*, IEEE, vol. 1, 1998, pp. 112–117.
- [57] K. R. Wheeler, D. A. Timucin, I. X. Twombly, K. F. Goebel, and P. F. Wysocki, “Aging aircraft wiring fault detection survey,” *NASA Ames Research Center*, 2007.
- [58] O. Osman, “Méthodes de diagnostic en ligne, embarqué et distribué dans les réseaux filaires complexes,” Ph.D. dissertation, Université Clermont Auvergne(2017-2020), 2020.
- [59] S. R. Skinner and K. Khan, “Wiring assessment of commuter class aircraft,” in *Aging Aircraft conference*, 2005.
- [60] J. G. Mahler, *Organizational learning at NASA: The Challenger and Columbia accidents*. Georgetown University Press, 2009.
- [61] M. Franchet, “Réflectométrie appliquée à la détection de défauts non francs dans les torons de câbles,” Ph.D. dissertation, Université Paris-Est, 2012.
- [62] F. AUZANNEAU, “Diagnostic filaire-détection, localisation et caractérisation de défauts dans des réseaux filaires complexes,” 2023.
- [63] C. R. Paul, *Analysis of multiconductor transmission lines*. John Wiley & Sons, 2007.
- [64] L. Sommervogel, “Various models for faults in transmission lines and their detection using time domain reflectometry,” *Progress In Electromagnetics Research C*, vol. 103, pp. 123–135, 2020.
- [65] P. Smith, C. Furse, and J. Gunther, “Analysis of spread spectrum time domain reflectometry for wire fault location,” *IEEE sensors journal*, vol. 5, no. 6, pp. 1469–1478, 2005.
- [66] V. Taylor and M. Faulkner, “Line monitoring and fault location using spread spectrum on power line carrier,” *IEE Proceedings-Generation, Transmission and Distribution*, vol. 143, no. 5, pp. 427–434, 1996.
- [67] A. Lelong and M. O. Carrion, “On line wire diagnosis using multicarrier time domain reflectometry for fault location,” in *SENSORS, 2009 IEEE*, IEEE, 2009, pp. 751–754.
- [68] T. Ajamian, “Exploration of compressive sampling for wire diagnosis systems based on reflectometry,” Ph.D. dissertation, École centrale de Nantes, 2019.
- [69] N. Ravot, “Analyse et diagnostic de réseaux filaires complexes par réflectométrie,” Ph.D. dissertation, Paris 11, 2007.

- [70] Y. C. Chung, C. Furse, and J. Pruitt, “Application of phase detection frequency domain reflectometry for locating faults in an f-18 flight control harness,” *IEEE transactions on Electromagnetic Compatibility*, vol. 47, no. 2, pp. 327–334, 2005.
- [71] H. M. Jol, *Ground penetrating radar theory and applications*. elsevier, 2008.
- [72] A. Benedetto and L. Pajewski, *Civil engineering applications of ground penetrating radar*. Springer, 2015.
- [73] M. Rasol, V. Pérez-Gracia, F. M. Fernandes, J. C. Pais, S. Santos-Assunção, and J. S. Roberts, “Ground penetrating radar system: principles,” *Handbook of Cultural Heritage Analysis*, pp. 705–738, 2022.
- [74] M. El-Said, “Geophysical prospection of underground water in the desert by means of electromagnetic interference fringes,” *Proceedings of the IRE*, vol. 44, no. 1, pp. 24–30, 1956.
- [75] M. Sato, H. Liu, T. Komagino, and K. Takahashi, “Archaeological survey by gpr for recovery from 3.11 great earthquake and tsunami in east japan,” in *Proceedings of the 15th International Conference on Ground Penetrating Radar*, IEEE, 2014, pp. 25–30.
- [76] M. Pipan, L. Baradello, E. Forte, and A. Prizzon, “Gpr study of bedding planes, fractures, and cavities in limestone,” in *Eighth International Conference on Ground Penetrating Radar*, SPIE, vol. 4084, 2000, pp. 682–687.
- [77] D. J. Daniels, “A review of gpr for landmine detection,” *Sensing and imaging*, vol. 7, no. 3, p. 90, 2006.
- [78] J. L. Davis and A. P. ANNAN, “Ground-penetrating radar for high-resolution mapping of soil and rock stratigraphy 1,” *Geophysical prospecting*, vol. 37, no. 5, pp. 531–551, 1989.
- [79] D. J. Daniels, “Ground penetrating radar for buried landmine and ied detection,” *Unexploded Ordnance Detection and Mitigation*, pp. 89–111, 2009.
- [80] A. Machado Brito-da-Costa, D. Martins, D. Rodrigues, L. Fernandes, R. Moura, and Á. Madureira-Carvalho, “Ground penetrating radar for buried explosive devices detection: a case studies review,” *Australian journal of forensic sciences*, vol. 54, no. 4, pp. 559–578, 2022.
- [81] M. Sow, V. Bertrand, N. Feix, M. Lalande, and E. Martinod, “Conception d’une antenne adaptée à un radar gpr impulsif pour une utilisation sur une tête de tunnelier design of antenna suitable for a gpr radar impulse use on a tunnel boring machine,”

- [82] R. G. Plumb, D. A. Noon, I. D. Longstaff, and G. F. Stickley, “A waveform-range performance diagram for ground-penetrating radar,” *Journal of applied Geophysics*, vol. 40, no. 1-3, pp. 117–126, 1998.
- [83] D. A. Noon, G. F. Stickley, and D. Longstaff, “A frequency-independent characterisation of gpr penetration and resolution performance,” *Journal of Applied Geophysics*, vol. 40, no. 1-3, pp. 127–137, 1998.
- [84] A. Lestari, A. Yarovoy, and L. Lighthart, “An efficient ultra-wideband bow-tie antenna,” in *2001 31st European Microwave Conference*, IEEE, 2001, pp. 1–4.
- [85] R. KING and T. WU, “The cylindrical antenna with non-reflecting resistive loading(distribution of current along center-driven cylindrical antenna with variable internal impedance per unit length),” 1964.
- [86] G. Dhatt, G. Touzot, and E. Lefrançois, *Méthode des éléments finis*. Lavoisier, 2005.
- [87] F. Hermeline, “Une méthode de volumes finis pour les équations elliptiques du second ordre,” *Comptes Rendus de l’Académie des Sciences-Series I-Mathematics*, vol. 326, no. 12, pp. 1433–1436, 1998.
- [88] D. DeCogan, *Transmission line matrix (TLM) techniques for diffusion applications*. CRC Press, 1998.
- [89] K. Yee, “Numerical solution of initial boundary value problems involving maxwell’s equations in isotropic media,” *IEEE Transactions on antennas and propagation*, vol. 14, no. 3, pp. 302–307, 1966.
- [90] A. Taflove, “Application of the finite-difference time-domain method to sinusoidal steady-state electromagnetic-penetration problems,” *IEEE Transactions on electromagnetic compatibility*, no. 3, pp. 191–202, 1980.
- [91] C. Guiffaut, “Contribution à la méthode ftd pour l’étude d’antennes et de la diffraction d’objets enfouis,” Ph.D. dissertation, Université de Rennes 1, école doctorale Matisse, 2000.
- [92] J. T. Hendrickson and P. J. Miller, “Progress in field testing of shielded cables,” in *[1992] Proceedings IEEE/AIAA 11th Digital Avionics Systems Conference*, IEEE, 1992, pp. 500–502.
- [93] C. Pavot and J. Liottin, “Câbles aéronautiques: les câbles spéciaux,” *REE. Revue de l’électricité et de l’électronique*, no. 6-7, pp. 55–56, 2003.
- [94] N. Taki, C. Delpha, D. Diallo, W. B. Hassen, and N. Ravot, “Soft fault diagnosis in wiring networks using reflectometry and principal component analysis,” *Measurement*, vol. 198, p. 111 378, 2022.

- [95] P. K. Kuhn, C. Furse, and P. Smith, “Locating hidden hazards in electrical wiring,” in *Aged Electrical Systems Research Application Symposium*, Citeseer, vol. 1, 2006, p. 8.
- [96] C. Jullien, P. Besnier, M. Dunand, and I. Junqua, “Analyse theorique et experimentale du couplage entre une paire torsadee non-blindee et un cable monofilaire,” in *EMC France*, 2012.
- [97] C. Furse, Y. C. Chung, R. Dangol, M. Nielsen, G. Mabey, and R. Woodward, “Frequency-domain reflectometry for on-board testing of aging aircraft wiring,” *IEEE Transactions on Electromagnetic Compatibility*, vol. 45, no. 2, pp. 306–315, 2003.
- [98] G. M. Gaeris and P. Z. Vanderlaan, *High-performance data cable*, US Patent 5,789,711, Aug. 1998.
- [99] M. K. Smail, “Développement d’une méthodologie dédiée à la réflectométrie en vue du diagnostic filaire,” Ph.D. dissertation, Paris 11, 2010.
- [100] H. M. Hashemian, “State-of-the-art predictive maintenance techniques,” *IEEE Transactions on Instrumentation and measurement*, vol. 60, no. 1, pp. 226–236, 2010.
- [101] E. J. Lundquist, J. R. Nagel, S. Wu, B. Jones, and C. Furse, “Advanced forward methods for complex wire fault modeling,” *IEEE Sensors Journal*, vol. 13, no. 4, pp. 1172–1179, 2012.
- [102] J. J. Han, S. R. Park, and S. Noh, “Improvement of detection performance in stdr/sstdr using sign detector,” *Electronics Letters*, vol. 51, no. 16, pp. 1281–1282, 2015.
- [103] C. Bishop and M. Tipping, “Pattern analysis and machine intelligence,” *IEEE Transactions on*, vol. 20, no. 3, pp. 281–293, 1998.
- [104] H. F. Senter *et al.*, “Applied linear statistical models . michael h. kutner, christopher j. nachtsheim, john neter, and william li,” *Journal of the American Statistical Association*, vol. 103, pp. 880–880, 2008.
- [105] U. Von Luxburg, “A tutorial on spectral clustering,” *Statistics and computing*, vol. 17, pp. 395–416, 2007.
- [106] H. Abdi, D. Valentin, B. Edelman, and A. J. O’Toole, “A widrow–hoff learning rule for a generalization of the linear auto-associator,” *Journal of Mathematical Psychology*, vol. 40, no. 2, pp. 175–182, 1996.
- [107] D. M. Clark and K. Ravishankar, “A convergence theorem for grossberg learning,” *Neural Networks*, vol. 3, no. 1, pp. 87–92, 1990.

- [108] H. Lu, Y. Li, M. Chen, H. Kim, and S. Serikawa, “Brain intelligence: go beyond artificial intelligence,” *Mobile Networks and Applications*, vol. 23, pp. 368–375, 2018.
- [109] Z. Yanling, D. Bimin, and W. Zhanrong, “Analysis and study of perceptron to solve xor problem,” in *The 2nd International Workshop on Autonomous Decentralized System, 2002.*, IEEE, 2002, pp. 168–173.
- [110] A. H. Dekker, “Kohonen neural networks for optimal colour quantization,” *Network: Computation in Neural Systems*, vol. 5, no. 3, p. 351, 1994.
- [111] J. A. Anderson, *An introduction to neural networks*. MIT press, 1995.
- [112] G. Joya, M. Atencia, and F. Sandoval, “Hopfield neural networks for optimization: study of the different dynamics,” *Neurocomputing*, vol. 43, no. 1-4, pp. 219–237, 2002.
- [113] A. M. Goh and X. L. Yann, “A novel sentiments analysis model using perceptron classifier,” *Int. J. of Electronics Engineering and Applications*, vol. 9, no. 4, pp. 01–10, 2021.
- [114] V. Dumoulin and F. Visin, “A guide to convolution arithmetic for deep learning,” *arXiv preprint arXiv:1603.07285*, 2016.
- [115] F. Yu and V. Koltun, “Multi-scale context aggregation by dilated convolutions,” *arXiv preprint arXiv:1511.07122*, 2015.
- [116] K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” *arXiv preprint arXiv:1409.1556*, 2014.
- [117] H. Bertrand, R. Ardon, M. Perrot, and I. Bloch, “Hyperparameter optimization of deep neural networks: combining hyperband with bayesian model selection,” in *Conférence sur l’Apprentissage Automatique*, 2017.
- [118] D. J. Daniels, “Surface-penetrating radar,” *Electronics & Communication Engineering Journal*, vol. 8, no. 4, pp. 165–182, 1996.
- [119] S. Lameri, F. Lombardi, P. Bestagini, M. Lualdi, and S. Tubaro, “Landmine detection from gpr data using convolutional neural networks,” in *2017 25th European Signal Processing Conference (EUSIPCO)*, IEEE, 2017, pp. 508–512.
- [120] D. A. Noon, “Stepped-frequency radar design and signal processing enhances ground penetrating radar performance,” 1996.
- [121] D. W. Paglioni, D. H. Chambers, J. E. Mast, S. W. Bond, and N. R. Beer, “Imaging modes for ground penetrating radar and their relation to detection performance,” *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 8, no. 3, pp. 1132–1144, 2014.

- [122] D. J. Daniels, “Ground penetrating radar 2nd edition,” *The Institution of Engineering and Technology, London*, 2007.