

Imagerie ultrasonore par réseaux antagonistes génératifs
conditionnels (cGAN)

par

NATHAN MOLINIER

MÉMOIRE PRÉSENTÉ À L'ÉCOLE DE TECHNOLOGIE SUPÉRIEURE
COMME EXIGENCE PARTIELLE À L'OBTENTION DE LA MAÎTRISE
AVEC MÉMOIRE EN GÉNIE MÉCANIQUE
M. Sc. A.

MONTREAL, LE 20 DÉCEMBRE 2022

ÉCOLE DE TECHNOLOGIE SUPÉRIEURE
UNIVERSITÉ DU QUÉBEC



Nathan Molinier, 2022



Cette licence [Creative Commons](#) signifie qu'il est permis de diffuser, d'imprimer ou de sauvegarder sur un autre support une partie ou la totalité de cette œuvre à condition de mentionner l'auteur, que ces utilisations soient faites à des fins non commerciales et que le contenu de l'œuvre n'ait pas été modifié.

PRÉSENTATION DU JURY

CE MEMOIRE A ETE EVALUÉ

PAR UN JURY COMPOSÉ DE :

M. Pierre Bélanger, directeur de mémoire
Département de génie mécanique à l'École de technologie supérieure

M. Matthew Toews, codirecteur de mémoire
Département de génie des systèmes à l'École de technologie supérieure

M. Ricardo J. Zednik, président du jury
Département de génie mécanique à l'École de technologie supérieure

M. Martin Viens, membre du jury
Département de génie mécanique à l'École de technologie supérieure

IL A FAIT L'OBJET D'UNE SOUTENANCE DEVANT JURY ET PUBLIC

LE 13 NOVEMBRE 2022

A L'ÉCOLE DE TECHNOLOGIE SUPERIEURE

REMERCIEMENTS

Je souhaite tout d'abord remercier mon directeur, le professeur Pierre Bélanger du département génie mécanique, pour m'avoir permis de réaliser ce projet et de l'avoir mené dans la direction qui me convenait le mieux. Je tiens également à le remercier pour son suivi et ses conseils avisés qui m'ont permis d'atteindre ces résultats mais aussi pour les opportunités de conférences et de publication qui se sont ouvertes à moi grâce à sa confiance.

Je souhaite également remercier mon co-directeur, le professeur Matthew Toews du département de génie des systèmes, pour son suivi et ses retours sur mon projet.

Je tiens ensuite à remercier Evident Industrial pour son financement me permettant ainsi de travailler dans de bonnes conditions pour ma maîtrise, mais aussi Alain Le Duff et Guillaume Painchaud-April pour leur suivi et leurs commentaires pertinents jusqu'au bout de mon projet.

Mes salutations respectueuses envers tous les membres du PulÉTS également, sans qui mon projet aurait été beaucoup plus difficile ! Je tiens particulièrement à remercier Baptiste Gauthier sans qui mon projet aurait mis beaucoup plus de temps à démarrer, autant pour ses cours particuliers d'imagerie ultrasonore que par son soutien informatique.

Merci également aux Arts et Métiers mon école d'ingénieur française et l'École de Technologie Supérieure de Montréal pour m'avoir proposé ce formidable cursus de double diplôme.

Merci enfin à ma famille pour m'avoir soutenu dans ce voyage et dans mon travail.

Imagerie ultrasonore par conditional generative adversarial network (cGAN)

NATHAN MOLINIER

RÉSUMÉ

La mise en forme des faisceaux avec la méthode de somme et de retard (de l'anglais Delay-And-Sum : DAS) en combinaison avec une acquisition de type capture matricielle complète (de l'anglais Full-Matrix-Capture : FMC) est une méthode encore aujourd'hui extrêmement utilisée dans le contrôle non destructif pour un grand nombre d'applications d'imagerie ultrasonore. Cependant, dans certains cas, comme les inspections à haute cadence, il est impossible d'envisager une telle méthode sans ralentir le processus. C'est pourquoi dans cette étude, nous avons cherché à contourner cette méthode standard en utilisant un réseau adversaire génératif conditionnel (de l'anglais conditional Generative Adversarial Network : cGAN) en combinaison avec des acquisitions type onde plane à zéro degré. Dans ce mémoire, trois différents modèles de cGAN ont été testés avec différentes fonctions de coût dans trois configurations différentes et les résultats ont été comparés à ceux obtenus grâce à la méthode conventionnelle. Les différents modèles ont été capables de reconstruire avec des résolutions équivalentes les images tout en améliorant dans plus de la moitié des cas le contraste, et en réduisant le niveau de bruit et les artefacts dans les images. De plus, la méthode proposée a permis de diviser le temps de calcul par 120 et le temps d'acquisition par 64 tout en divisant l'espace de stockage nécessaire aux données d'entrée par 75.

Mots clés : DAS, TFM, FMC, Contrôle non destructif, Ultrasons, Onde plane (PW), cGAN

Imagerie ultrasonore par conditional generative adversarial network (cGAN)

NATHAN MOLINIER

ABSTRACT

The Full Matrix Capture (FMC) combined with the Total Focusing Method (TFM) are often considered as the gold standard in ultrasonic nondestructive evaluation. However, this method is not always convenient because of the amount of data required and the time required to gather the data. Indeed, when it comes to high cadence inspections, gathering the FMC and processing it may take too long. This study proposes to replace conventional FMC acquisition and TFM processing with a single zero-degree Plane Wave (PW) insonification and a conditional Generative Adversarial Networks (cGAN) trained to produce TFM-like images. Three models with different architecture and loss formulations were tested in different scenarios and their performances were compared with TFM images. The proposed cGAN were able to recreate the important features of the images and also improve the contrast in more than half the reconstructions in comparison with conventional TFM reconstructions. The contrast was systematically increased through a reduction of the background noise level and the elimination of some artifacts. Finally, the proposed method led to a reduction of the computation time and file size by respectively a factor of 120 and 75.

Mots clés : DAS, TFM, FMC, Non destructive evaluation, Ultrasound, Plane wave (PW), cGAN

TABLE DES MATIÈRES

	Page
INTRODUCTION	21
CHAPITRE 1 REVUE DE LITTÉRATURE	23
1.1 Propagation des ondes ultrasonores	23
1.1.1 Formulation des équations 1D	23
1.1.2 Formulation des équations 3D	24
1.1.3 Comportement des ondes aux interfaces.....	27
1.1.4 Atténuation des ondes	29
1.2 Imagerie ultrasonore	29
1.2.1 Les sondes ultrasonores	30
1.2.2 Émission et réception mono-élément (A-Scan)	30
1.2.3 Émission et réception multiéléments	31
1.2.3.1 Émission en ondes plane	31
1.2.3.2 Émission avancée : Full Matrix Capture (FMC).....	33
1.2.4 Reconstruction d'images : Total Focusing Method (TFM)	33
1.2.4.1 Acquisition type FMC.....	34
1.2.4.2 Acquisition type ondes planes	35
1.3 Réseaux de neurones.....	35
1.3.1 Perceptrons et réseaux de neurones	35
1.3.2 Principe d'apprentissage	37
1.3.3 Introduction aux réseaux de neurones convolutifs.....	38
1.3.3.1 Les architectures	39
1.3.3.2 L'apprentissage	41
1.3.3.3 Le conditional Generative Adversarial Network (cGAN)	42
1.3.3.4 Exemples d'applications de l'intelligence artificielle à l'imagerie ultrasonore	43
1.4 Création d'une base de données.....	44
1.4.1 Simulation par éléments finis.....	44
1.4.1.1 Équation Élastodynamique discrète	45
1.4.1.2 Utilisation de frontières absorbantes.....	45
1.4.1.3 Dimension des éléments et incrément temporels	46
1.4.2 Création d'une base de données plus « réaliste »	46
1.4.2.1 Ajout de données expérimentales	46
1.4.2.2 Ajout d'un bruit expérimental.....	47
1.5 Conclusion	47
CHAPITRE 2 OBJECTIFS ET MÉTHODOLOGIE	49
2.1 Objectif	49
2.2 Méthodologie	50
2.2.1 Bases de données	50
2.2.1.1 Les images type.....	50
2.2.1.2 Création de données	51

2.2.2	Réseaux de neurones et architectures	55
2.2.2.1	Modèle discriminateur	56
2.2.2.2	Modèles générateurs	56
2.2.3	Fonctions de coût	59
2.2.4	Processus d'entraînement.....	61
2.2.5	Métriques	62
2.3	Conclusion	65
CHAPITRE 3 RESULTATS ET VALIDATION EXPERIMENTALE		67
3.1	Résultats.....	67
3.1.1	Reconstruction d'images.....	67
3.1.2	Contraste de signal sur bruit (CNR), erreur de position et API.....	72
3.1.3	Vitesse des modèles et taille des données.....	75
3.2	Discussion préliminaire	75
3.3	Discussion et validation des résultats	76
3.4	Conclusion et validation des résultats.....	79
CONCLUSION.....		81
RECOMMANDATIONS		83
BIBLIOGRAPHIE		85

LISTE DES TABLEAUX

	Page
Tableau 2.1	Plage des données simulées 52
Tableau 2.2	Propriétés du matériau simulé 53
Tableau 2.3	Propriété de la sonde utilisées 53
Tableau 2.4	Plage des données expérimentales 54
Tableau 2.5	Répartition des données dans les bases de données 55
Tableau 2.6	Modèles de cGAN utilisés 62
Tableau 3.1	Valeurs caractéristiques pour la différence de CNR, l'erreur de position et l'API 75
Tableau 3.2	Comparaison des performances entre le DAS et la méthode alternative utilisant un cGAN 75

LISTE DES FIGURES

	Page	
Figure 1.1	Illustration des deux modes de propagation ondulatoire adapté de Le Manuel numérique max - Belin Education	27
Figure 1.2	Schéma de propagation d'une onde à une interface	28
Figure 1.3	Schéma d'acquisition d'un A-Scan	31
Figure 1.4	Génération d'une onde plane par une loi de retards adapté de Dupont- Marilla, (2018)	32
Figure 1.5	Schéma d'acquisition d'une FMC	33
Figure 1.6	Schéma d'une méthode de focalisation en tous points	34
Figure 1.7	Modèle du perceptron	36
Figure 1.8	Schéma d'un réseau de neurones avec 2 couches cachées	37
Figure 1.9	Exemples d'application CNN adapté de Mathew (2022)	38
Figure 1.10	Représentation des cartes de caractéristiques d'un modèle de CNN Adapté de B. Li & Xu (2015)	39
Figure 1.11	Architecture de VGG16 Extrait de Le (2021)	40
Figure 1.12	Transformations d'image lésion de la peau Adapté de Mikołajczyk & Grochowski (2018)	42
Figure 2.1	Image TFM et image d'une onde plane d'un SDH dans un block d'acier inoxydable	51
Figure 2.2	Schémas des paramètres utilisés pour les deux types de défauts	52
Figure 2.3	Modèle discriminateur	56
Figure 2.4	Modèle générateur 1	57
Figure 2.5	Modèle générateur 2	59
Figure 2.6	Processus d'entraînement	61
Figure 3.1	Reconstruction d'un même SDH L'onde plane d'entrée est dans la plage d'entraînement des réseaux	68

Figure 3.2	Reconstruction d'un même FBH L'onde plane d'entrée est dans la plage d'entraînement des réseaux	69
Figure 3.3	Reconstruction d'un même SDH L'onde plane d'entrée est en dehors de la plage d'entraînement des réseaux le matériau est différent.....	70
Figure 3.4	Reconstruction de deux SDH L'onde plane d'entrée est en dehors de la plage d'entraînement des réseaux le matériau est différent et le nombre de défaut est différent	71
Figure 3.5	Différence de CNR entre la TFM générée et la TFM de référence pour les données testées (a) dans la plage de l'entraînement et (b) en dehors de la plage de l'entraînement	72
Figure 3.6	Erreur de position du défaut dans la TFM image générée pour les données testées (a) dans la plage de l'entraînement et (b) en dehors de la plage de l'entraînement La position du défaut dans la TFM de référence est considérée exacte.....	73
Figure 3.7	Différence d'API entre la TFM image générée et la TFM de référence pour les données testées (a) dans la plage de l'entraînement et (b) en dehors de la plage de l'entraînement	74

LISTE DES ABRÉVIATIONS, SIGLES ET ACRONYMES

CND	Contrôle Non Destructif
FBH	Trou à fond plat (traduction de Flat Bottom Hole)
SDH	Trou percé latéral (traduction de Side Drilled Hole)
ROI	Région d'intérêt (traduction de Region Of Interest)
GPU	Processeur graphique (traduction de Graphics Processing Unit)
3D	Tridimensionnel
2D	Bidimensionnel
SNR	Ratio signal sur bruit (traduction de Signal to Noise Ratio)
TFM	Méthode de focalisation en tout point (traduction de Total Focusing Method)
FMC	Capture matricielle complète (traduction de Full Matrix Capture)
DAS	Somme et retard (traduction de Delay And Sum)
ReLU	Unité rectifiée linéaire (traduction de Rectified Linear Unit)
CNN	Réseau neuronal convolutif (traduction de Convolutional Neural Network)
FC	Entièrement connecté (traduction de Fully Connected)
cGAN	Réseaux antagonistes génératifs conditionnels (traduction de conditional Generative Adversarial Network)
G	Générateur
D	Discriminateur
FE	Éléments finis (traduction de Finite Element)
EMAT	Sonde ultrasonore à électro-aimant (traduction de Electro-Magnetic Acoustic Transducer)
API	Interface d'application programmable (traduction de Application Programming Interface)
BDD	Base de données
RVB	Rouge Vert Bleu
VGG	Groupe Géométrique Visuel (traduction de Visual Geometry Group)
CNR	Contraste de signal sur bruit (traduction de Contrast to Noise Ratio)

LISTE DES SYMBOLES ET UNITÉS DE MESURE

dB	Décibel
Hz	Hertz
m	Mètre
s	Seconde
m/s	Mètre par seconde
E	Module d'Young
ν	Coefficient de Poisson
λ	Longueur d'onde
c	Vitesse de l'onde
ρ	Masse volumique
σ	Fonction d'activation
w	Poids des réseaux de neurones
b	Biais du perceptron
s	Sortie des réseaux de neurones
x	Entrée des réseaux de neurones
y	Image de référence de la base de données
B	Octet (traduction de Byte)

INTRODUCTION

L'imagerie ultrasonore par sonde multiéléments est un procédé fortement utilisé dans le cadre du contrôle non destructif (CND). En effet, grâce à sa forte adaptativité et sa facilité d'utilisation, ce procédé s'est démarqué dans de nombreuses applications, comme l'inspection de soudures, l'analyse d'épaisseur, la détection de fissures, l'inspection de corrosions... L'objectif de ce procédé est de venir reconstruire une image de la section d'une pièce à l'aide de la propagation d'ondes ultrasonores afin d'assurer sa viabilité dans différents contextes industriels comme l'aéronautique, l'industrie automobile, l'industrie pétrolière...

De nombreux algorithmes ont pu voir le jour afin d'améliorer ces méthodes d'imagerie. L'objectif étant toujours de développer des images possédant la meilleure résolution possible le plus rapidement possible. Cependant, l'inconvénient de la plupart des méthodes d'imagerie actuelles est l'importante quantité de données nécessaire à leur exécution. Ce problème est double, puisque plus de données implique un plus long temps d'acquisition mais également un plus long temps de traitement, ce qui devient limitant pour diminuer le temps de reconstruction.

Pour pallier ces problèmes, cette étude propose une méthode de traitement assistée par réseaux de neurones. L'objectif de cette méthode est tout d'abord de venir limiter la quantité d'émission de la sonde à une onde plane afin de réduire le temps d'acquisition et d'utiliser un réseau type conditional Generative Adversarial Network (cGAN) pour apprendre à reconstruire des images de plus haute résolution à partir de cette acquisition de faible résolution. Ce rapport est construit en deux parties, un premier chapitre de revue de littérature décrivant les concepts de base nécessaires à la compréhension du second chapitre constitué d'un article décrivant la méthode énoncée précédemment et présentant des résultats de comparaison entre les différentes méthodes d'imagerie.

CHAPITRE 1

REVUE DE LITTÉRATURE

1.1 Propagation des ondes ultrasonores

Les ondes ultrasonores appartiennent à la famille des ondes mécaniques. Elles ne peuvent donc se propager que dans des milieux dit non-vides comme les gaz, les liquides ou encore les solides. Leurs déplacements se manifestent par des variations de pression ou de contraintes au sein des différents milieux et leurs fréquences vibratoires sont situées au-delà du domaine audible de 20kHz. Différentes ondes ultrasonores peuvent être définies en fonction de leur direction de propagation et de leur direction de vibration.

1.1.1 Formulation des équations 1D

L'équation de propagation des ondes dans un milieu unidimensionnel, isotrope et homogène a été définie par d'Alembert à partir des lois de Hooke et de Newton ainsi que de l'hypothèse permettant d'assimiler les liaisons entre les particules d'un solide par un système masse-ressort (Cheeke, 2002). L'équation de propagation est définie par :

$$\frac{\partial^2 u(x, t)}{\partial t^2} = c^2 \frac{\partial^2 u(x, t)}{\partial x^2} \quad (1.1)$$

Avec u représentant le déplacement de la particule en fonction de la position x et du temps t , et c la vitesse de propagation. La solution générale à l'équation (1.1) peut s'exprimer comme suit :

$$u(x, t) = f(x - c \cdot t) + g(x + c \cdot t) \quad (1.2)$$

Avec f et g deux fonctions de propagation arbitraires respectivement dans le sens des x positifs et des x négatifs. Si on considère une perturbation sinusoïdale dans le sens des x positifs, l'équation devient :

$$u(x, t) = A \cos \left(\omega \left(t - \frac{x}{c} \right) \right) \quad (1.3)$$

Avec A l'amplitude de l'onde et ω sa pulsation angulaire.

1.1.2 Formulation des équations 3D

L'équation de propagation des ondes a ensuite été généralisée par Euler pour les cas tridimensionnels (3D) solides, homogènes et isotropes sans force extérieure appliquée (Shull, 2002) :

$$\rho \frac{\partial^2 u_i}{\partial t^2} = \frac{\partial T_{ij}}{\partial x_j} \quad (1.4)$$

Avec ρ la masse volumique du matériau de propagation, et T_{ij} son tenseur de contrainte obtenu à l'aide de la loi de Hooke :

$$T_{ij} = c_{ijkl} \cdot S_{kl} \quad (1.5)$$

Avec C le tenseur d'élasticité de coefficients c_{ijkl} et S le tenseur des déformations de coefficients S_{kl} . T_{ij} peut également s'écrire :

$$T_{ij} = \lambda S_{ii} \delta_{ij} + 2\mu S_{ij} \quad (1.6)$$

Avec λ et μ les paramètres de Lamé que l'on peut décrire à partir du module d'Young (E) et du coefficient de Poisson (ν) ou des coefficients c_{ijkl} du tenseur d'élasticité :

$$\left\{ \begin{array}{l} E = \mu \frac{3\lambda + 2\mu}{\lambda + \mu} \\ \nu = \frac{\lambda}{2(\lambda + \mu)} \\ c_{44} = \frac{c_{11} - c_{12}}{2} = \mu \\ c_{12} = \lambda \end{array} \right. \quad (1.7)$$

On réécrit l'expression (1.6) en utilisant (1.7) :

$$T_{ij} = c_{12}S_{ii}\delta_{ij} + 2c_{44}S_{ij} = (c_{11} - 2c_{44})\frac{\partial u_i}{\partial x_i}\delta_{ij} + 2c_{44} \times \frac{1}{2} \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right) \quad (1.8)$$

Finalement, si on se ramène à l'expression (1.4) on obtient :

$$\rho \frac{\partial^2 u_i}{\partial t^2} = \frac{\partial}{\partial x_i} \left[(c_{11} - c_{44}) \frac{\partial u_j}{\partial x_i} \right] + c_{44} \frac{\partial^2 u_i}{\partial x_j^2} \quad (1.9)$$

Sous sa forme vectorielle :

$$\rho \frac{\partial^2 \vec{u}}{\partial t^2} = (c_{11} - c_{44}) \vec{\nabla}(\vec{\nabla} \cdot \vec{u}) + c_{44} \Delta \vec{u} \quad (1.9)$$

Si on utilise la décomposition de Helmholtz, le vecteur \vec{u} se découpe en un champ vectoriel $\vec{\Psi}$ et un champ scalaire Φ :

$$\vec{u} = \Phi \vec{\nabla} + \vec{\nabla} \times \vec{\Psi} = \vec{u}_L + \vec{u}_T \quad (1.10)$$

Avec \vec{u}_L la composante longitudinale de l'onde vibrant normalement par rapport au plan de propagation, et \vec{u}_T la composante transversale de l'onde vibrant dans le plan de propagation. En combinant les équations (1.9) et (1.10) il est possible d'isoler ces deux composantes :

$$\begin{cases} \frac{\partial^2 \Phi}{\partial t^2} - c_L^2 \nabla^2 \Phi = 0 \\ \frac{\partial^2 \vec{\Psi}}{\partial t^2} - c_T^2 \nabla^2 \Psi = 0 \end{cases} \quad (1.11)$$

Il est donc possible d'en déduire les vitesses de propagation des ondes longitudinales et transversales :

$$\begin{cases} c_L = \sqrt{\frac{\lambda + 2\mu}{\rho}} \\ c_T = \sqrt{\frac{\mu}{\rho}} \end{cases} \quad (1.12)$$

Les ondes longitudinales et transversales sont donc les seules ondes se propageant dans les milieux solides, homogènes et isotropes. De ce fait toutes les ondes peuvent être décrites comme des combinaisons linéaires de ces deux modes de propagation. La Figure 1.1 illustre ces deux modes de propagation.

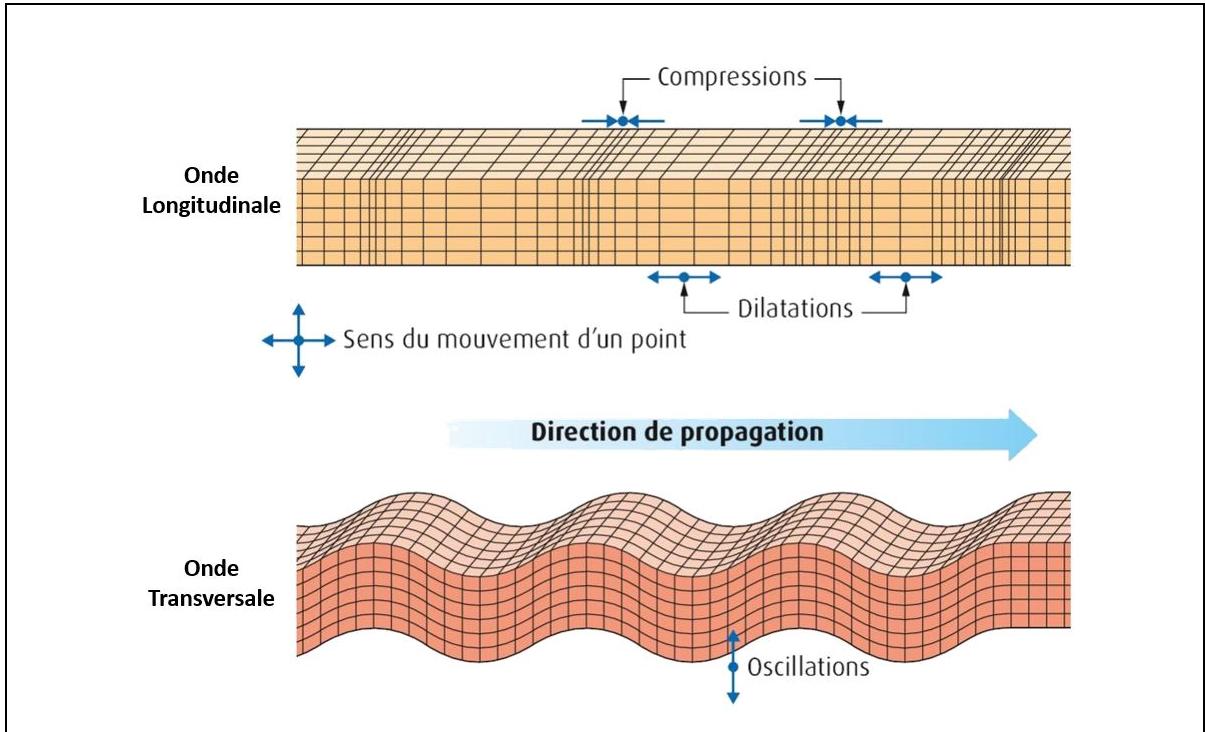


Figure 1.1 Illustration des deux modes de propagation ondulatoire
adapté de Le Manuel numérique max - Belin Education

1.1.3 Comportement des ondes aux interfaces

Lorsqu'une onde se propage, elle peut être amenée à rencontrer différents milieux (solide, liquide, gaz). De ce fait, en fonction des propriétés de ces différents milieux ainsi que de l'angle d'incidence de l'onde par rapport à cette interface, plusieurs phénomènes physiques sont amenés à se produire. Une partie de l'énergie est transmise tandis que l'autre est réfléchie. Cette interaction s'appelle la réfraction et celle-ci est définie par la loi de Snell-Descartes :

$$\frac{\sin \theta_i}{c_i} = \frac{\sin \theta_{tL}}{c_{tL}} = \frac{\sin \theta_{tT}}{c_{tT}} = \frac{\sin \theta_{rL}}{c_{rL}} = \frac{\sin \theta_{rT}}{c_{rT}} \quad (1.13)$$

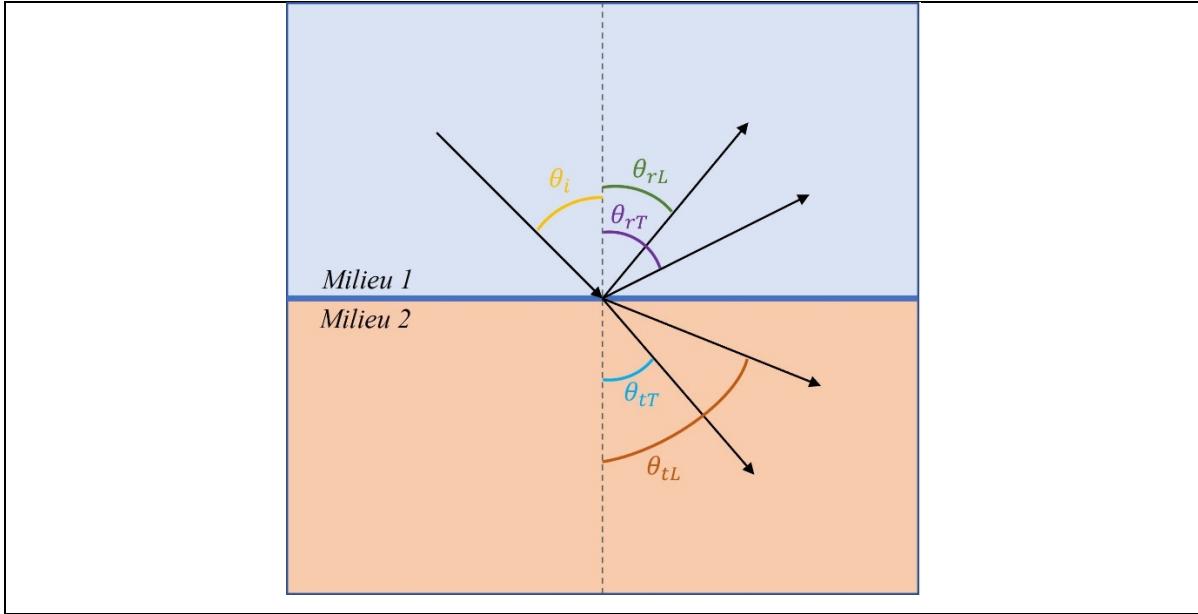


Figure 1.2 Schéma de propagation d'une onde à une interface

Avec θ_i l'angle d'incidence de l'onde, θ_{rL} l'angle de l'onde longitudinale réfléchie, θ_{rT} l'angle de l'onde transversale réfléchie, θ_{tL} l'angle de l'onde longitudinale transmise et θ_{tT} l'angle de l'onde transversale transmise. Les paramètres c_i , c_{rL} , c_{rT} , c_{tL} et c_{tT} désignent respectivement les vitesses de ces ondes.

Afin de quantifier de l'énergie transmise et réfléchie par une onde à l'interface entre deux milieux, il est possible de définir également les coefficients de réflexion (R) et de transmission (T) à partir des impédances acoustiques des deux milieux. Dans le cas d'une incidence normale, ces coefficients s'expriment de la manière suivante :

$$\begin{cases} R = \frac{Z_2 - Z_1}{Z_2 + Z_1} \\ T = \sqrt{1 - R^2} \end{cases} \quad (1.14)$$

Avec $Z_n = \rho_n \cdot c_n$ l'impédance acoustique du milieu n .

1.1.4 Atténuation des ondes

Lorsqu'une onde se propage dans un milieu, son énergie est amenée à se dissiper. Ce phénomène est appelé atténuation et se manifeste par une décroissance exponentielle de son amplitude :

$$A(x) = A_0 \cdot e^{-\alpha x} \quad (1.15)$$

Avec A_0 l'amplitude initiale de l'onde, x la distance parcourue et α le coefficient d'atténuation. Ce dernier est influencé par de nombreux paramètres généralement divisés en deux catégories, la diffusion du milieu liée aux qualités microscopiques et macroscopiques des matériaux, la présence ou non de fissures, de dislocations, de joints de grains... et l'absorption qui est liée aux effets thermiques où ici l'atténuation est proportionnelle au carré de la fréquence de l'onde (Cheeke, 2002).

Une solution pour pallier cette atténuation lors d'une acquisition pourrait être d'ajouter un gain équivalent permettant de compenser la diminution d'amplitude du signal, néanmoins une telle correction ne tiendrait pas compte d'un autre problème très présent lors de n'importe quelle acquisition expérimentale : le bruit. En effet, qu'il soit d'origine électronique ou dû à une non-homogénéité parfaite du matériau, le bruit est dans certains cas difficile à isoler du signal pur. Le SNR (ratio de signal sur bruit traduction de Signal to Noise Ratio) est communément utilisé afin de quantifier de cette différence d'amplitude entre le signal et le bruit en décibel (dB). La mise en place de certains filtres est également possible afin d'augmenter le SNR en diminuant le niveau de bruit.

1.2 Imagerie ultrasonore

Le concept d'imagerie ultrasonore, dans le contrôle non destructif (CND), n'est pas récent et fait partie du paysage industriel depuis plusieurs années. En effet, ce procédé est très utilisé dans de nombreux domaines comme l'aéronautique, l'industrie automobile, l'industrie pétrolière... pour venir réaliser des inspections toujours de plus en plus précises. De ce fait, de

nombreuses méthodes d'imagerie ont pu voir le jour, et encore aujourd'hui les enjeux restent importants. L'objectif étant toujours d'améliorer la résolution des images tout en augmentant la cadence d'inspection.

1.2.1 Les sondes ultrasonores

Pour réaliser des émissions ultrasonores, de nombreuses technologies peuvent être utilisées : les sondes piézoélectriques, les EMAT (Electro-Magnetic Acoustic Transducer) ... Dans notre cas, nous nous concentrerons sur l'utilisation de sondes piézoélectriques car plus simples à utiliser dans le cadre de l'imagerie. Leur aspect et leur fonction peut varier suivant plusieurs paramètres : le nombre d'éléments piézoélectriques, l'agencement des éléments (mono-élément, linéaires, matricielles...), l'angle d'incidence de l'onde (wedge, contact...), la température d'utilisation, le milieu d'utilisation (air, eau, ...), ...

1.2.2 Émission et réception mono-élément (A-Scan)

Pour comprendre les concepts de l'imagerie par ultrasons, il faut tout d'abord comprendre le type d'acquisition réalisable à l'aide d'un élément piézoélectrique.

Le A-Scan constitue l'information la plus brute que l'on peut obtenir avec un élément piézoélectrique. Celui-ci représente l'amplitude ondulatoire de l'élément au cours du temps. Cette information permet de mesurer assez simplement la distance entre deux points de réflexion de l'onde incidente (réflexions sur des interfaces différentes).

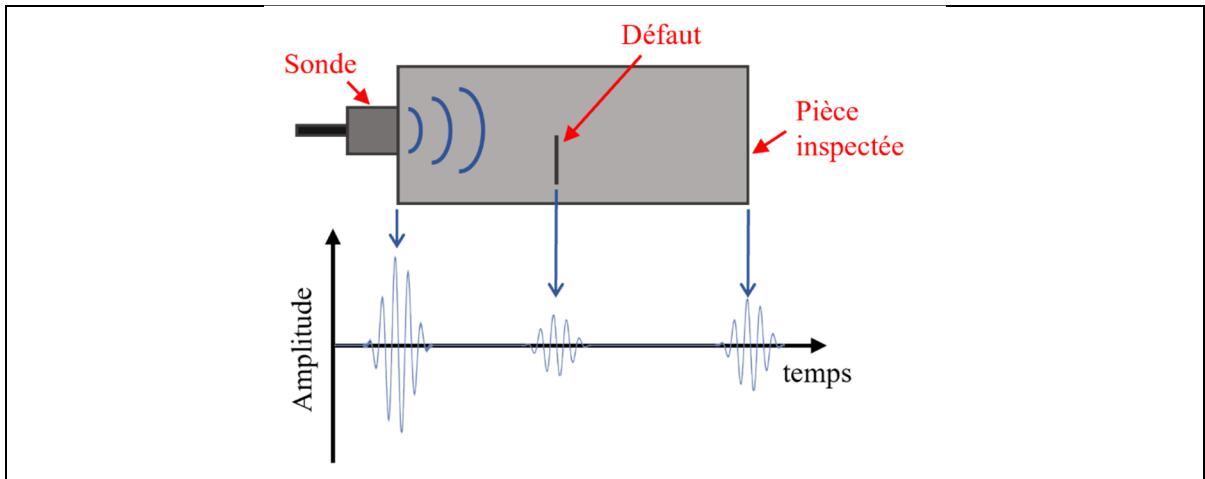


Figure 1.3 Schéma d'acquisition d'un A-Scan

La Figure 1.3 représente un schéma d'acquisition d'un A-Scan pour une sonde en contact, le premier écho correspond à l'émission du signal, le deuxième à la réflexion de l'onde sur le défaut et le troisième à la réflexion de l'onde sur le fond de la pièce.

1.2.3 Émission et réception multiéléments

Lorsqu'on parle d'imagerie ultrasonore, l'objectif est de venir recréer la section d'une pièce grâce à la propagation des ultrasons. Les sondes multiéléments sont alors couramment utilisées car plus larges, plus puissantes et plus versatiles du fait de leurs multiples combinaisons d'émission/réception. Dans cette partie, nous aborderons deux méthodes d'émission/réception.

1.2.3.1 Émission en ondes plane

Cette méthode fortement utilisée dans le domaine médical (échographie) est une des méthodes proposant les meilleurs résultats en terme de vitesse d'acquisition (Couture et al., 2012). Cependant, à cause du manque de focalisation de cette dernière, la qualité de l'image créée se retrouve limitée en termes de résolution (Montaldo et al., 2009). Cette méthode consiste à venir générer des fronts d'ondes avec tous les éléments de la sonde.

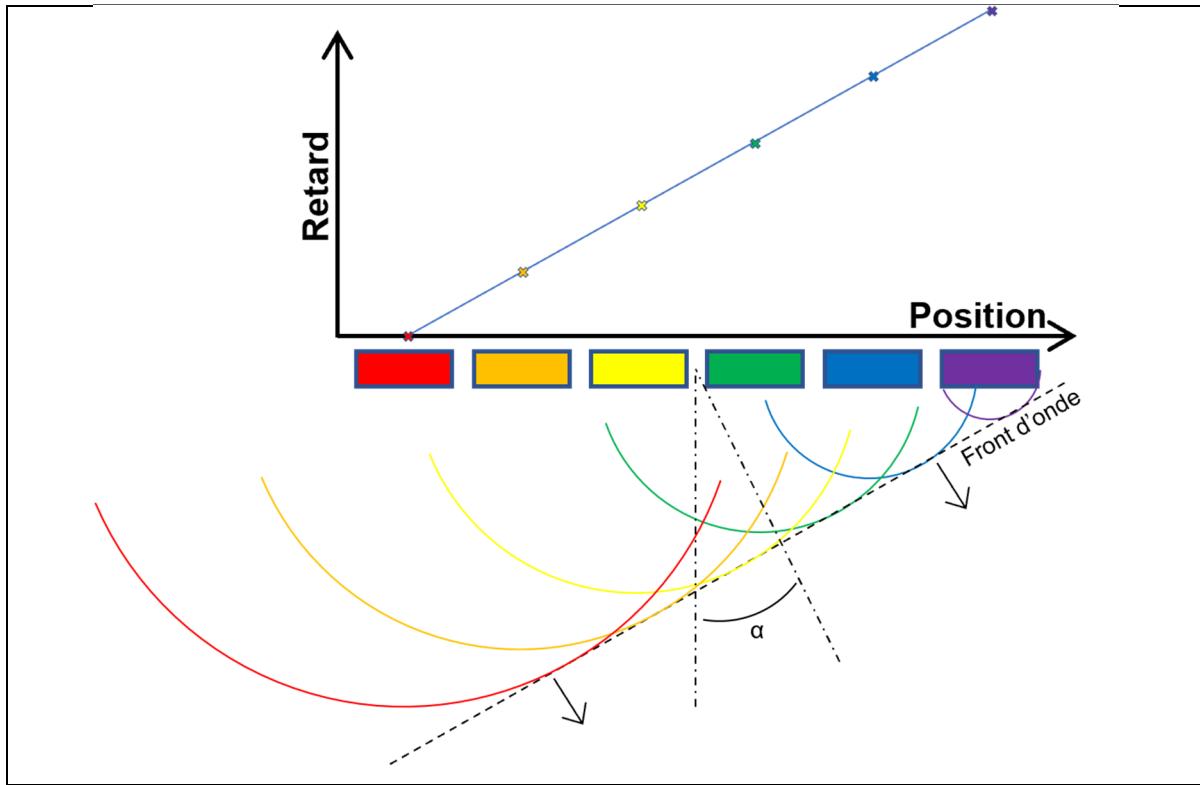


Figure 1.4 Génération d'une onde plane par une loi de retards
adapté de Dupont-Marilla, (2018)

La Figure 1.4 présente ce type d'émission pour un angle d'émission α . La loi de retard entre deux éléments peut être définie comme suivant :

$$\Delta t = \frac{p}{c} \sin \alpha \quad (1.16)$$

Avec p la distance entre deux éléments, c la vitesse de propagation de l'onde, et Δt le retard entre deux éléments successifs.

Remarque : Lors de ce type d'émission, tous les éléments restent actifs pour la réception des échos pendant tout le temps d'acquisition.

Dans le chapitre suivant, une onde plane à 0-degré sera utilisée, ce qui implique un tir simultané de tous les éléments et donc un retard $\Delta t = 0$.

1.2.3.2 Émission avancée : Full Matrix Capture (FMC)

Pour pallier aux problèmes de focalisation rencontrés avec la méthode précédente, une acquisition type FMC est régulièrement mise en œuvre (Schmerr, 2015). Cette méthode consiste à émettre successivement avec chacun des éléments d'une sonde multiéléments tout en recevant simultanément sur tous les éléments. La Figure 1.5 illustre cette méthode.

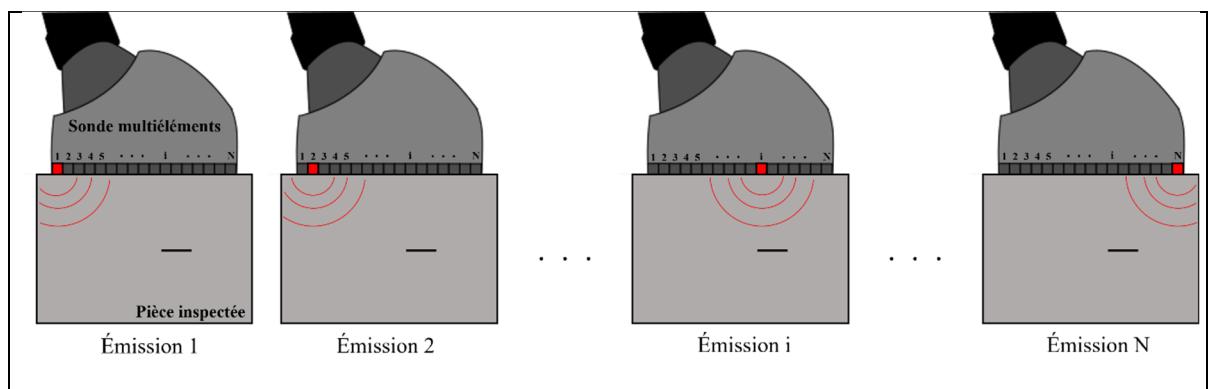


Figure 1.5 Schéma d'acquisition d'une FMC

Remarque : tous les éléments y compris celui qui émet sont actifs pour la réception des échos

Les défauts d'une telle méthode sont le temps d'acquisition qui est beaucoup plus important que pour une émission type onde plane (environ N fois plus long), et la taille de la matrice obtenue qui est pour une sonde de N éléments, N fois plus importante que pour une acquisition type onde plane.

1.2.4 Reconstruction d'images : Total Focusing Method (TFM)

Dans la partie précédente, nous avons pu voir plusieurs méthodes d'émissions/réceptions ultrasonores à l'aide de sondes multiéléments. Cependant, un post-traitement est nécessaire afin de reconstruire une image de la section de la pièce. Il existe plusieurs types de méthodes communément appelées « Total Focusing Method » (TFM) comme le Delay-And-Sum

beamforming (DAS) (Perrot et al., 2021), ou encore le Minimum Variance Beamforming (MVB) (Synnevag et al., 2009)…

1.2.4.1 Acquisition type FMC

L'objectif de ces méthodes est simple, calculer les délais des différents échos grâce à la vitesse de l'onde pour chaque pixel de la région d'intérêt (ROI) à l'aide des A-Scans de la FMC (quadrillage blanc sous la sonde Figure 1.6),

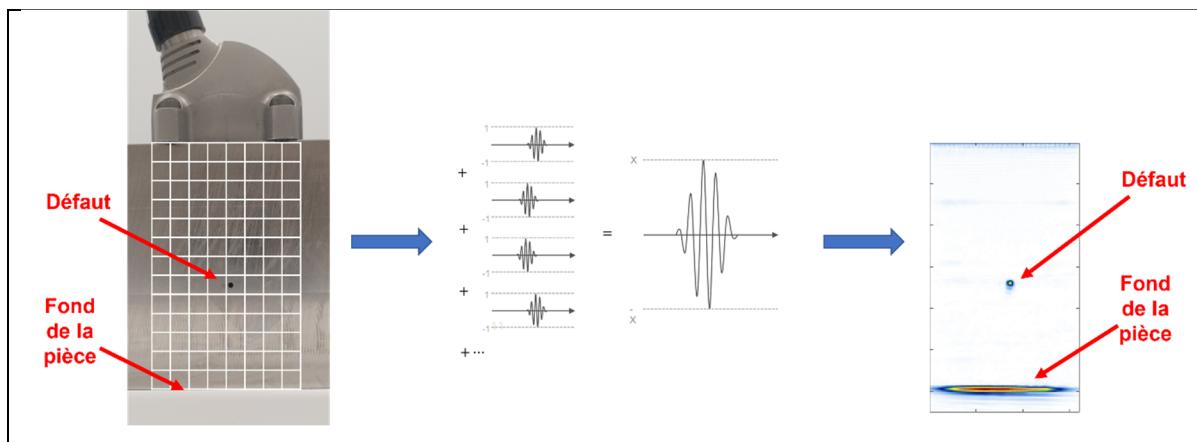


Figure 1.6 Schéma d'une méthode de focalisation en tous points

Puis sommer les amplitudes de ces signaux en tenant compte de la position de chaque élément piézoélectrique de la sonde. Les délais sont préalablement appliqués à ces signaux afin de les superposer de manière cohérente (interférences constructives), ce qui permet de reconstruire une image de la pièce, à droite sur la Figure 1.6. La formule de somme peut être définie comme suivant :

$$I_{TFM}(x, z) = \left| \sum_{k=1}^N \sum_{l=1}^N |A_{kl}(x, z)| e^{i\phi_{kl}(x, z)} \right| \quad (1.17)$$

Avec $I_{TFM}(x, z)$ la valeur d'un pixel reconstruit de coordonnées (x, z) , $A_{kl}(x, z)$ l'amplitude d'un signal et $\phi_{kl}(x, z)$ sa phase et N le nombre d'éléments de la sonde.

Remarque : Même si la résolution de l'image proposée par cette méthode est importante, cette méthode reste lente et peut conduire à de faibles contrastes ou encore à la génération d'artéfacts (Asl & Mahloojifar, 2009).

1.2.4.2 Acquisition type ondes planes

Dans le cadre de cette acquisition, les méthodes de TFM peuvent également être utilisées, et en fonction du nombre d'ondes planes et d'angles utilisés, il est possible d'améliorer la résolution de reconstruction. Cette méthode permet un bon compromis entre vitesse d'acquisition et résolution d'image.

Cependant, dans cette étude, une seule émission à zéro-degré a été réalisée afin d'optimiser au maximum le temps d'acquisition.

1.3 Réseaux de neurones

Le concept de réseaux de neurones n'est pas récent, en effet, les premiers modèles ont pu voir le jour dans les années 1940. Cependant, ce n'est que depuis les années 2000 que ceux-ci sont réellement considérés grâce aux progrès importants de l'informatique et notamment depuis la création des premiers GPU (Graphic Processing Unit) facilitant le calcul parallèle. Ce domaine est également un axe grandissant dans la recherche actuelle car les applications sont nombreuses et ce dans de nombreux domaines.

1.3.1 Perceptrons et réseaux de neurones

Tout d'abord, avant de parler de réseau de neurones, il faut comprendre ce que constitue une unité élémentaire : un « neurone ». Celui-ci, plus communément appelé perceptron ou réseau monocouche a été conceptualisé par Frank Rosenblatt en 1958 à partir des travaux réalisés par McCulloch & Pitts dans les années 1940. Il s'agit d'un opérateur mathématique à plusieurs entrées x_k mais à sortie unique s comme le montre la formule suivante :

$$s = \sigma \left(\sum_{k=1}^n w_k x_k + b \right) \quad (1.20)$$

Avec w_k appelés poids donnant plus ou moins d'importance à chacune des entrées x_k , b appelé biais et σ une fonction d'activation. Le premier modèle de fonction d'activation utilisé était la fonction Heaviside, définie comme suivant :

$$\sigma_{Heaviside}(x) = \begin{cases} 1 & \text{si } x > 0 \\ 0 & \text{sinon} \end{cases} \quad (1.21)$$

Depuis de nouvelles fonctions d'activation sont apparues et aujourd'hui, l'une des fonctions d'activation les plus utilisées est la fonction ReLU (Rectified Linear Unit) présentée ci-dessous :

$$\sigma_{ReLU}(x) = \begin{cases} x & \text{si } x > 0 \\ 0 & \text{sinon} \end{cases} \quad (1.22)$$

La Figure 1.7 présente le schéma du concept de perceptron.

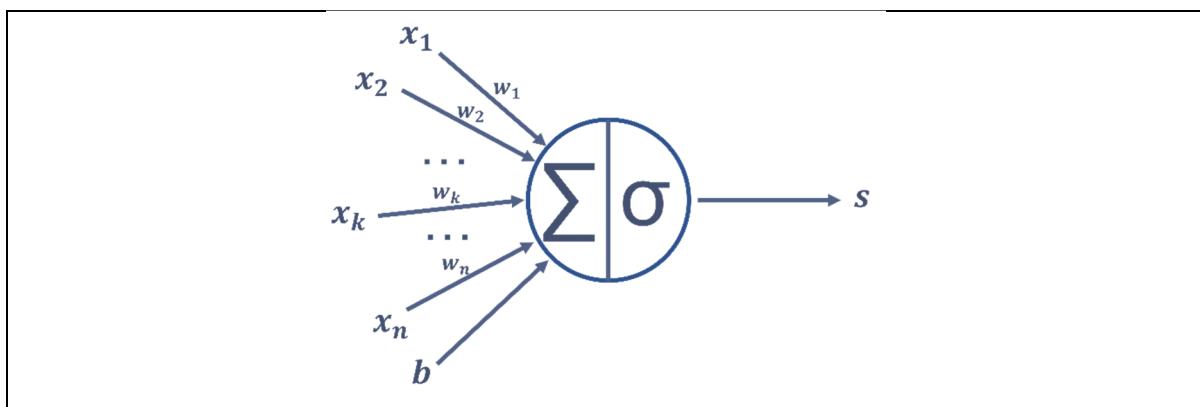


Figure 1.7 Modèle du perceptron

Concernant les réseaux de neurones, ceux-ci sont en fait constitués de multiples perceptrons interconnectés. La Figure 1.8 illustre ce principe avec un réseau de neurones interconnectés.

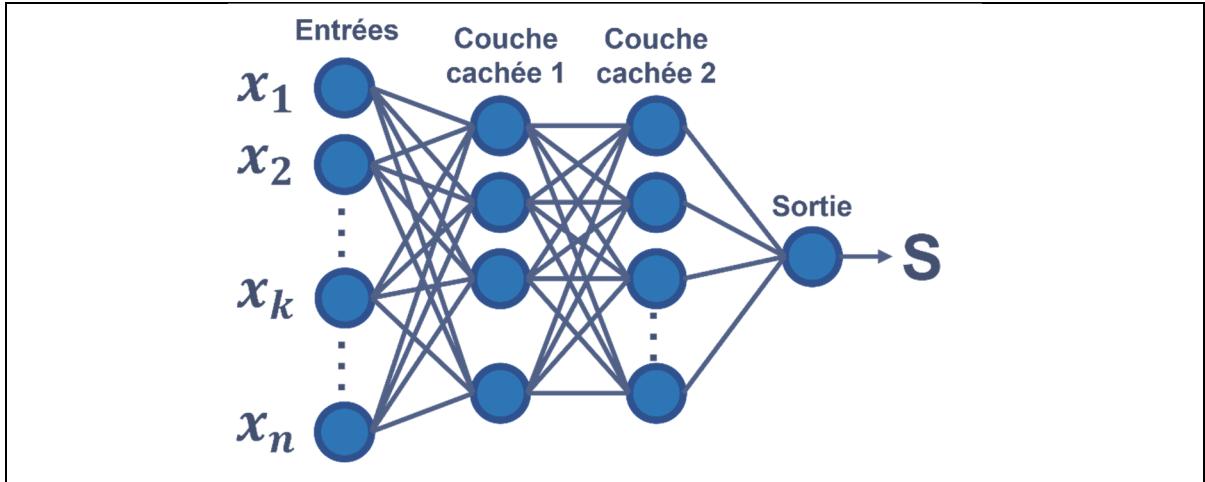


Figure 1.8 Schéma d'un réseau de neurones avec 2 couches cachées

Remarque : il est plus rare d'observer des réseaux complètement interconnectés car plus lourds à entraîner.

1.3.2 Principe d'apprentissage

L'objectif dans l'apprentissage machine, est de venir développer une capacité à pouvoir généraliser un comportement, une situation. Cet apprentissage peut être vu comme l'optimisation d'une fonction complexe où les paramètres sont les poids w_k qui constituent la mémoire du réseau. On cherche à minimiser l'écart entre la sortie réelle s du réseau et la sortie attendue y à l'aide de plusieurs itérations à partir de données. Cet écart est aussi appelé fonction de coût C . La fonction de coût la plus commune est l'erreur quadratique moyenne :

$$C = \frac{1}{N} \sum_{k=1}^N (s_i(w) - y_i)^2 \quad (1.23)$$

Pour minimiser cette fonction de coût C nous avons recours à une méthode appelée la descente de gradient :

$$\frac{\partial C}{\partial t} = \frac{\partial C}{\partial w} \frac{\partial w}{\partial t} \leq 0 \quad (1.24)$$

À chaque étape d'entraînement, les poids du réseau sont ajustés par rétropropagation et les couches sont ainsi mises à jour.

Remarque : Pour pallier l'arrivée des plus gros modèles et donc des plus larges bases de données d'entraînement, le gradient stochastique est apparu : la minimisation de la fonction de coût est plus rapide car l'ajustement des poids se fait après chaque exemple et non après le passage de toute la base de données. Le défaut d'une telle méthode est de venir trouver un minimum local et non global dû à une moins bonne généralisation des données.

1.3.3 Introduction aux réseaux de neurones convolutifs

Les réseaux de neurones convolutifs ou CNN (Convolutional Neural Networks) sont des architectures multicouches utilisées principalement dans le cadre du traitement d'images. Les applications les plus communes sont la classification d'images, la segmentation d'instances, la détection, ... La Figure 1.9 présente la différence entre ces différents types d'application.

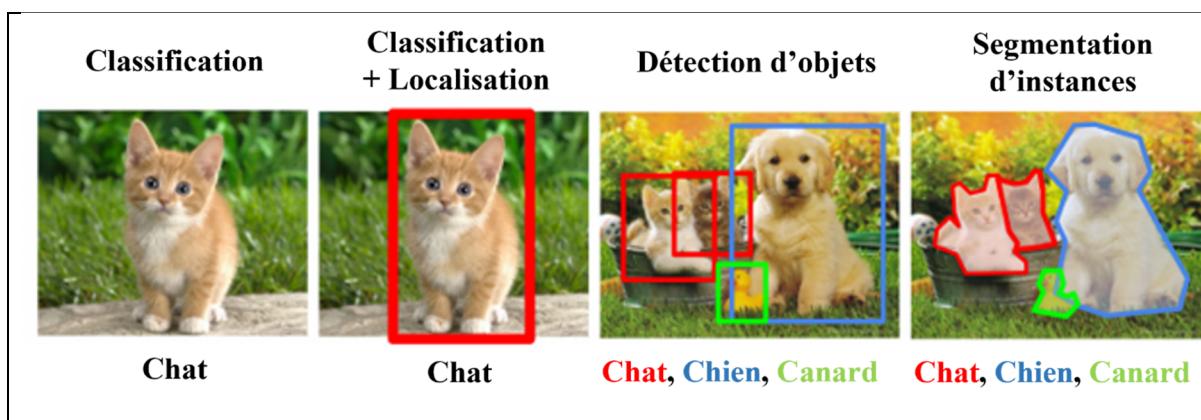


Figure 1.9 Exemples d'application CNN
adapté de Mathew (2022)

1.3.3.1 Les architectures

Les CNN tentent d'imiter par leur fonctionnement les mécanismes de vision des hommes et des animaux. Ce principe se rapproche fortement des filtres de Sobel (Kanopoulos et al., 1988) utilisés en imagerie dans le cadre de la détection de contours.

Dans le cadre de la vision par ordinateur, pour une application donnée, l'objectif de l'entraînement d'un CNN est de venir créer des filtres successifs qui seront utilisés afin d'extraire des caractéristiques des images en entrée. Ces filtres plus communément appelés feature map (« cartes de caractéristiques ») vont venir se concentrer chacun sur des caractéristiques différentes des images comme l'illustre la Figure 1.10.

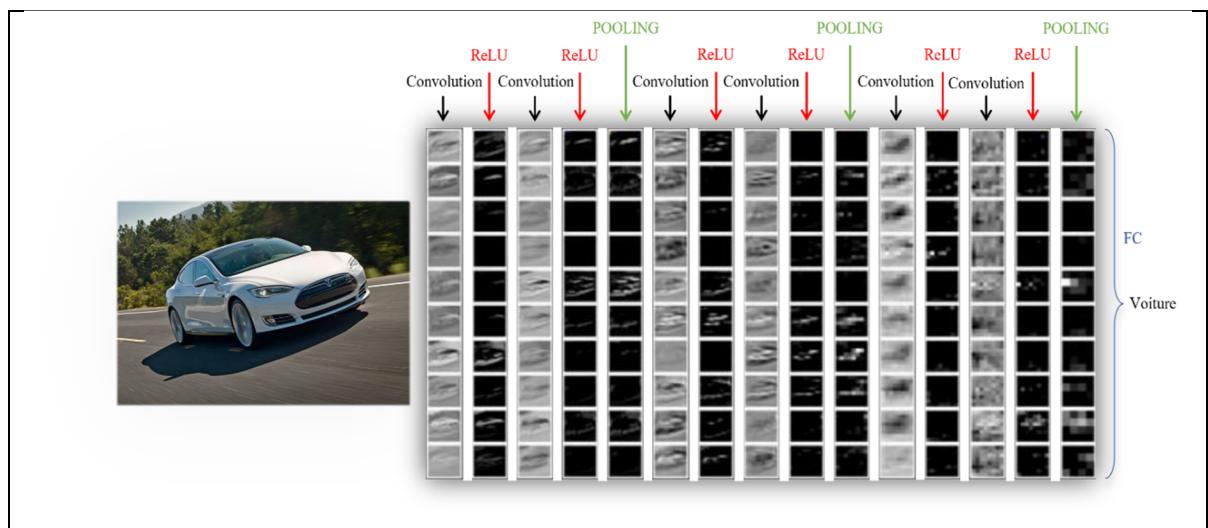


Figure 1.10 Représentation des cartes de caractéristiques d'un modèle de CNN
Adapté de B. Li & Xu (2015)

La Figure 1.10 présente également les 4 types d'étape que l'on retrouve dans la plupart des réseaux à convolutions :

- Les couches de convolution : couches possédant un ou plusieurs filtres responsables de l'extraction des caractéristiques des images. Pour créer ces filtres, on fait glisser sur

l'image une matrice (« kernel ») et on réalise un produit matriciel (produit de Hadamard) pour chaque position de cette matrice sur l'image.

- Les couches de pooling : couches permettant de réduire le nombre de paramètres que le réseau doit apprendre et donc les risques de surapprentissage.
- Les fonctions d'activation (généralement ReLU) : Pour améliorer la vitesse de l'entraînement en supprimant les valeurs négatives mais en conservant les valeurs positives.
- Les couches de neurones interconnectées (FC : Fully Connected) : Pour se ramener à un vecteur de classification, la probabilité pour chacune des sorties (ici une voiture et non un camion, un avion, un bateau ou un cheval)

La Figure 1.11 présente un exemple d'architecture de CNN, ici VGG16.

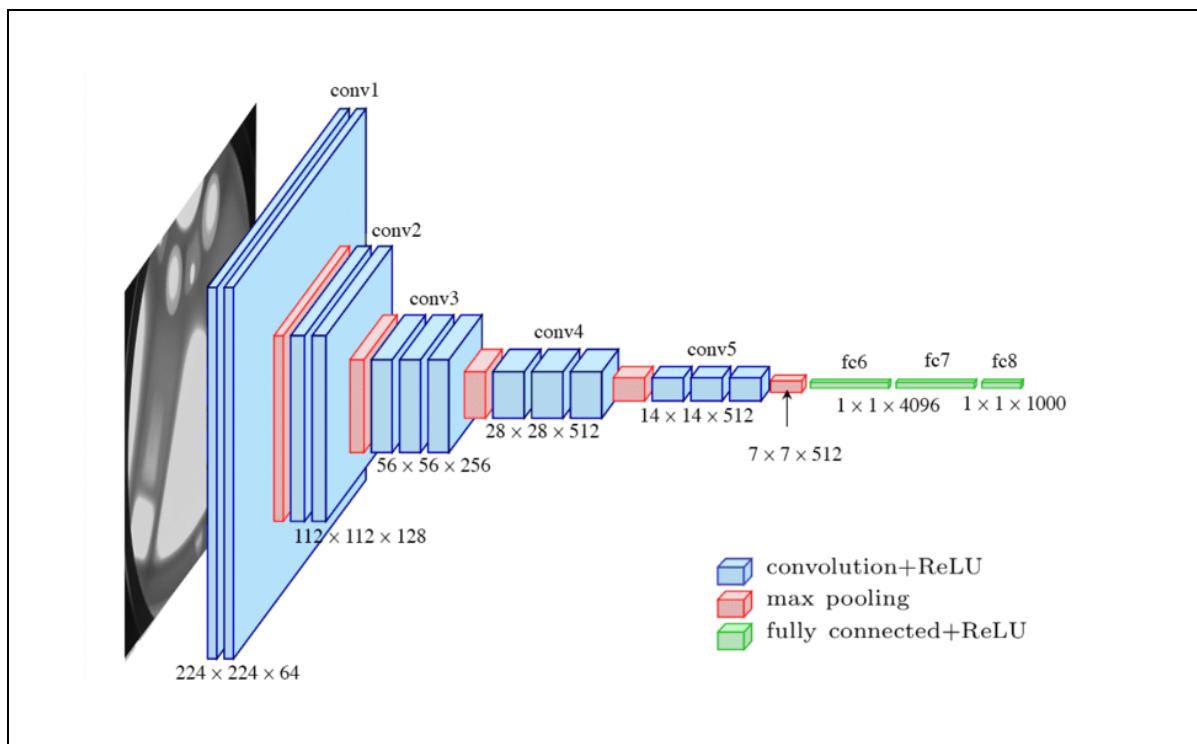


Figure 1.11 Architecture de VGG16
Extrait de Le (2021)

1.3.3.2 L'apprentissage

Au cours de l'apprentissage des CNN ou modèles profonds, plusieurs images sont utilisées afin d'entraîner ses capacités de vision dans un contexte donné. Néanmoins, en fonction de la base de données utilisée, les conditions d'entraînement peuvent être amenées à varier considérablement.

Tout d'abord, il est important de considérer la taille de la base de données. En effet, si la base de données est trop petite ou trop inégale en termes de données (par exemple, pour un problème de classification, présence d'une quantité inégale de données correspondant à chacune des classes à identifier), il y a un risque de mauvais apprentissage voire de surapprentissage (ou overfitting en anglais) du réseau car impossibilité de ce dernier à pouvoir généraliser.

Ensuite, posséder ou non des données étiquetées (données avec des labels, des détourages dans l'image...) est un deuxième enjeu. Dans le cadre d'un entraînement disposant de 100% des données étiquetées, on parle d'entraînement supervisé, mais ce genre d'apprentissage est généralement plus coûteux et plus difficile à obtenir. C'est pourquoi d'autres types d'entraînement ont été introduits, comme l'apprentissage semi-supervisé (Weston et al., 2008) où une majorité des données ne sont pas étiquetées ou encore l'apprentissage dit « renforcé » où les données ne sont pas étiquetées complètement : par exemple les classes sont définies mais non localisées. On parle également d'apprentissage faiblement supervisé (Durand, 2017) et d'apprentissage non-supervisé, dans ces cas, l'objectif du réseau est de venir rassembler des similarités entre les images grâce à des techniques de regroupement afin d'en distinguer des classes.

Pour pallier ces problèmes, il existe de nombreuses techniques d'augmentation de données (Mikołajczyk & Grochowski, 2018). Une des techniques la plus réputée et simple à mettre en place est la transformation de données déjà étiquetées. On vient appliquer des transformations aléatoires type rotation, symétrie, distorsion, recadrage... aux images afin de proposer au réseau un autre point de vue sur une donnée déjà connue et ainsi améliorer la capacité de

généralisation de ce dernier ainsi que sa robustesse. Il est aussi possible d'ajouter des filtres de couleur pour modifier la donnée comme le présente la Figure 1.12.

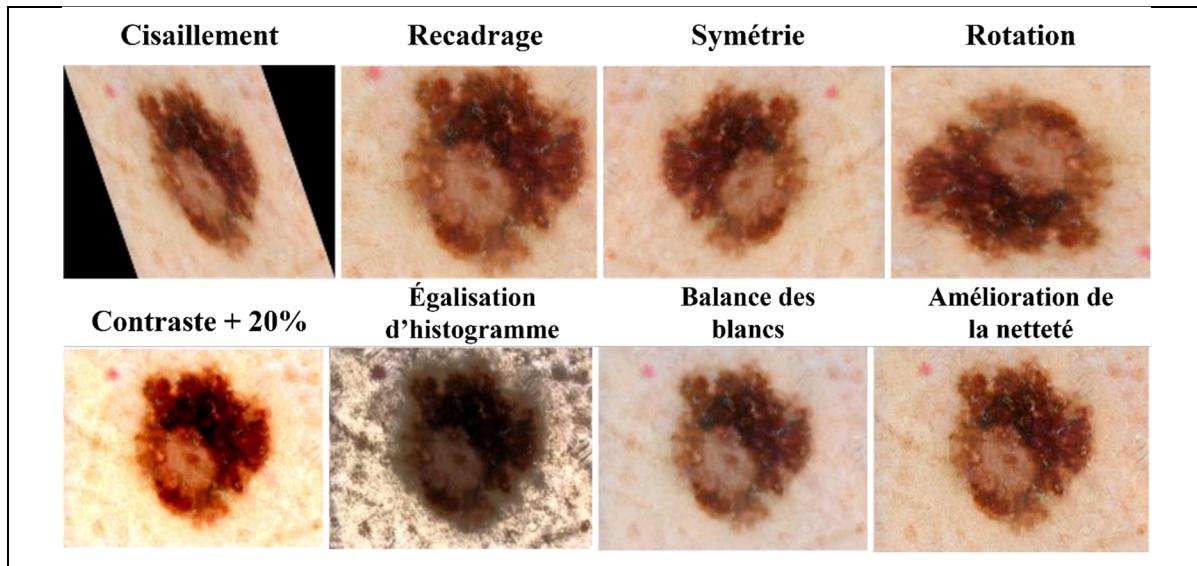


Figure 1.12 Transformations d'image lésion de la peau
Adapté de Mikołajczyk & Grochowski (2018)

1.3.3.3 Le conditional Generative Adversarial Network (cGAN)

Une autre application courante des modèles profonds est la création d'images. Dans ce domaine, de nombreuses applications ont pu voir le jour. Des algorithmes de super-résolution où l'objectif est de venir reconstruire une image avec une meilleure résolution que l'image initiale (Tang et al., 2021), ou encore des méthodes visant à améliorer la qualité des images en temps réel (Qi et al., 2021)...

C'est dans ce contexte que certains modèles ont pu se détacher du lot comme les cGAN.

L'objectif du cGAN est d'apprendre à reconstruire une image y label à partir d'une image x en entrée (Isola et al., 2018). Cette tâche revient alors à apprendre au réseau une répartition non-linéaire entre ces deux images.

Pour atteindre cet objectif, le réseau est en fait décomposé en deux sous réseaux possédant des tâches très différentes :

- Un modèle **générateur G** dont l'objectif est de venir construire à partir d'une entrée x , une image y quasi identique à celle provenant de la base de données.
- Un modèle **discriminateur D** dont l'objectif est de distinguer les images « réelles » de la base de données avec celles générées par le modèle générateur « fausses ».

La raison pour laquelle ce réseau s'est démarqué est sa capacité à pouvoir obtenir des résultats sans nécessiter beaucoup d'optimisation. En effet, le concept de réseaux antagonistes est efficace puisque le modèle discriminateur fonctionne en fait comme une fonction de coût qui va venir s'optimiser seule pendant l'entraînement.

Cette fonction de coût peut être décrite comme suivant :

$$L_{adv} = E_{x,y}[\log D(x, y)] + E_{x,y}[\log(1 - D(x, G(x, z)))] \quad (1.25)$$

Avec x l'image d'entrée, y l'image label et z un vecteur de bruit aléatoire. G essaie de minimiser cet objectif pendant que D essaie de le maximiser.

1.3.3.4 Exemples d'applications de l'intelligence artificielle à l'imagerie ultrasonore

Les applications liées à l'imagerie ultrasonores sont nombreuses et sont présentes dans de nombreux contextes (industriel, médical...)

Ces applications sont liées à :

- de la classification d'images (Burgos-Artizzu et al., 2020; Q. Li et al., 2014; Munir et al., 2019; Sai Sundar et al., 2019). Munir et al. (2019) propose l'utilisation d'un CNN permettant la classification de défauts au sein de soudures,
- de la détection d'anomalies (Y.-C. Li et al., 2021; Virupakshappa et al., 2018; Yap et al., 2018). Virupakshappa et al., (2018) propose lui aussi l'utilisation d'un CNN pour venir localiser des fissures au sein d'images ultrasonores.
- de la segmentation d'images (Nair et al., 2018; Zhou et al., 2019; Zhuang et al., 2019). Zhou et al. (2019) propose une méthode de segmentation par CNN visant à analyser

l'apparition de plaques d'athérome pour prévenir les risques d'Athérosclérose de la carotide.

- de la reconstruction d'images (Pavlov et al., 2019; Perdios et al., 2022; Wang et al., 2020). Wang et al. (2020) propose l'utilisation d'un cGAN pour venir reconstruire des images de haute résolution à partir de donnée brute de faible résolution.

D'après Litjens et al., (2017) l'intelligence artificielle est maintenant devenue un incontournable dans la recherche liée à l'imagerie médicale.

1.4 Création d'une base de données

Comme nous l'avons vu dans la partie précédente, posséder une base de données de qualité est essentiel pour entraîner un réseau de neurones. L'objectif de cette partie est d'introduire les différentes méthodes utilisées afin de venir générer des données.

1.4.1 Simulation par éléments finis

L'un des moyens les plus communs pour venir générer des données est le recours aux simulations par éléments finis. Les logiciels les plus utilisés dans le cadre des inspections ultrasonores sont Abaqus, Pogo, Ansys...

Dans notre cas, nous nous sommes intéressé à Pogo (Huthwaite, 2014). Grâce à son utilisation de processeurs graphiques (GPU : Graphique Processing Unit), Pogo est capable de paralléliser les calculs sur plusieurs coeurs ce qui diminue fortement le temps nécessaire à la simulation. Et son API (Application Programming Interface) compatible avec Matlab, facilite également la modification de n'importe quel paramètre de simulation. Ces caractéristiques rendent Pogo extrêmement pertinent et rapide pour tester facilement différentes configurations d'inspections impliquant une multitude de paramètres.

1.4.1.1 Équation Élastodynamique discrète

L'équation élastodynamique discrétisée dans le domaine spatial se formule comme suit :

$$M \frac{d^2\vec{u}}{dt^2} + C \frac{d\vec{u}}{dt} + K\vec{u} = \vec{F} \quad (1.18)$$

Avec \vec{u} le vecteur de déplacement, M la matrice diagonale de masse dont les valeurs dépendent de la masse volumique ρ , K la matrice de rigidité statique dont les valeurs dépendent du module d'Young E et du coefficient de Poisson ν , C la matrice d'amortissement visqueux dépendant de l'amortissement de Rayleigh et \vec{F} la force extérieure appliquée. L'état d'équilibre est perturbé par l'application d'une force qui entraîne la propagation d'ondes dans le maillage des éléments finis.

Pour résoudre cette équation, les positions de chaque nœuds sont calculées à chaque incrément temporel : les détails de résolution de cette équation sont visibles dans l'article de Huthwaite, (2014). Les types d'éléments peuvent être de plusieurs natures : 3D (éléments tétraédriques ou cubiques) ou encore 2D (éléments triangulaires ou carrés) en fonction du nombre de dimensions du modèle. L'intérêt d'une modélisation 2D plutôt que 3D est le gain en termes de vitesse de calcul. En effet dans de nombreux cas le gain en complexité d'un modèle 3D n'apporte pas de réel changement lors de la simulation ce qui pousse généralement à formuler l'hypothèse que le modèle en 2D est équivalent.

1.4.1.2 Utilisation de frontières absorbantes

Afin de réduire la complexité des modèles simulés, il est courant d'utiliser des frontières absorbantes dans les modèles. Par exemple, si on souhaite diminuer l'impact des réflexions latérales sur les bords de la pièce pendant la simulation, il est possible d'augmenter la taille du modèle afin de négliger ces réflexions grâce à l'atténuation. Cependant, augmenter la taille du modèle augmente aussi la complexité de calcul et donc le temps de simulation. C'est pourquoi, afin de pallier ce problème, il est possible d'incorporer au modèle des frontières absorbantes,

dont le rôle est d'atténuer plus rapidement toute onde se propageant dans ce milieu. Ces régions sont en fait composées de plusieurs couches absorbantes successives, avec une gradation fine du coefficient d'absorption pour minimiser les potentielles réflexions.

1.4.1.3 Dimension des éléments et incrément temporels

Deux critères essentiels se doivent d'être respectés afin d'éviter toute erreur de simulation.

La taille des éléments doit être inférieure à la longueur d'onde utilisée pendant la simulation pour ne pas venir détecter d'autres modes d'excitation. Drozdz et al., (2006) recommandent un minimum de 15 éléments par longueur d'onde.

Et la taille d'un incrément temporel doit être inférieure au temps nécessaire à l'onde pour traverser un élément (Cook et al., 2007). Cette condition peut être exprimée comme suit :

$$\Delta t = 0.8 \frac{\Delta e_{min}}{c_{max}} \quad (1.19)$$

Avec Δe_{min} la taille de l'élément le plus petit, c_{max} la vitesse de l'onde la plus rapide et 0.8 un facteur de sécurité.

1.4.2 Crédation d'une base de données plus « réaliste »

Afin d'utiliser un réseau dans des conditions réelles d'imagerie, plusieurs méthodes vont être utilisées afin de complexifier la base de données.

1.4.2.1 Ajout de données expérimentales

Tout d'abord, des données réelles ont été ajoutées. Ces acquisitions ont été réalisées à l'aide d'un système Verasonics Vantage 64 LE ainsi qu'une sonde multiéléments Olympus de

référence 5L64-32X10-A32-P-2.5-OM. L'objectif de ces données est de venir améliorer les performances du réseau grâce à des conditions réelles d'utilisation.

1.4.2.2 Ajout d'un bruit expérimental

Afin de faire ressembler les images simulées à des images plus « expérimentales », un bruit a été ajouté à ces dernières afin d'imiter le bruit d'expérimentation (instruments, conditions de mesures...). Ce bruit a été généré aléatoirement de manière à suivre une distribution type Gaussienne. Il a été filtré dans la bande passante de la sonde utilisée expérimentalement et son amplitude a été déterminée de manière à respecter un SNR (ratio signal sur bruit) équivalent à celui observé dans les données expérimentales. Ce dernier a ensuite été ajouté individuellement sur chacun des A-Scans des FMC simulées ainsi que sur les données brutes correspondant aux acquisitions type onde plane.

1.5 Conclusion

Dans cette première partie, nous avons pu aborder plusieurs concepts fondamentaux nécessaires à la bonne compréhension de la suite du rapport. Une introduction au contrôle non destructif par ultrasons a été réalisée jusqu'aux concepts liés à l'imagerie ultrasonore, puis les concepts d'intelligence artificielle ont ensuite été abordés avant de terminer sur certaines méthodes clés mises en place dans le cadre de la conception d'une base de données. Enfin, une dernière partie présentant l'objectif du mémoire ainsi qu'une brève description de la méthodologie mise en pratique a été réalisée.

CHAPITRE 2

OBJECTIFS ET MÉTHODOLOGIE

Dans le chapitre précédent, nous avons pu aborder grâce à une revue de littérature les notions d'imagerie ultrasonore et d'intelligence artificielle. Ce chapitre vise à développer ces concepts dans notre cas d'application et à en préciser les objectifs.

2.1 Objectif

L'objectif de ce mémoire est de venir présenter une nouvelle méthode d'imagerie ultrasonore pertinente dans le cadre d'une application industrielle haute cadence.

Cette méthode se voudra :

- Rapide : autant en termes d'acquisition qu'en termes de vitesse de traitement des données
- Légère : en ressources et en données, on veut limiter l'espace de stockage
- Efficace : pour proposer des images de bonne résolution dans toutes les situations

Pour atteindre cet objectif :

- Une base de données comportant deux types de défauts (SDH et FBH) simulés et expérimentaux a été créée,
- Différentes architectures de cGAN et fonctions de coût provenant de la littérature ont été sélectionnées puis combinées en trois différents modèles,
- Ces modèles ont ensuite été entraînés puis testés dans des conditions similaires d'utilisation
- Différentes métriques ont été mises en place afin d'évaluer leurs performances et les comparer aux méthodes usuelles de reconstruction d'image.

2.2 Méthodologie

Dans cette partie seront abordés, les bases de données créées, les architectures et fonctions de coût utilisées par les modèles de réseaux de neurones, l'entraînement ainsi que les métriques utilisées pour comparer les performances des méthodes.

2.2.1 Bases de données

La création d'une base de données présente de nombreux enjeux comme vu précédemment. L'objectif de cette partie est de venir présenter la base de données qui a été utilisée ainsi que les choix qui ont été fait lors de sa création.

2.2.1.1 Les images type

Les images utilisées dans les bases de données sont de deux types :

- Des TFM créée à partir d'acquisitions type FMC (cf. p34)
- Des images créées à partir d'une onde plane à zéro-degré (cf. p35)

La Figure 2.1 présente l'allure de ces deux images pour un défaut type trou percé latéralement (traduction de side-drilled hole : SDH).

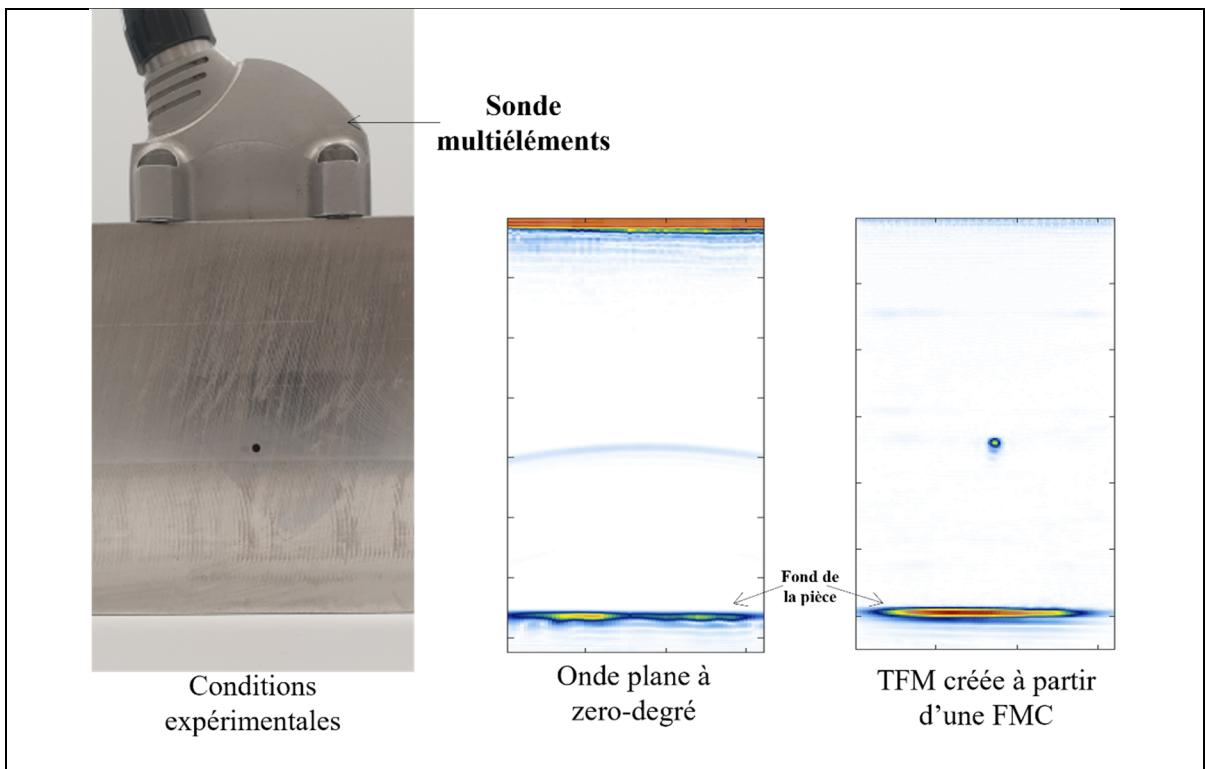


Figure 2.1 Image TFM et image d'une onde plane d'un SDH dans un block d'acier inoxydable

2.2.1.2 Crédit de données

Dans cette étude, une base de données composées de données simulées et expérimentales a été créée afin d'entraîner différents modèles de réseaux de neurones. Les données simulées ont été utilisées afin d'augmenter la taille de la base de données et de diminuer la nécessité de collecter des données expérimentales qui sont très difficiles à obtenir car très coûteuses. Ensuite, des défauts conventionnellement utilisés dans des blocks de calibration ont servi lors de la création des données : des trous percés latéralement (traduction de side-drilled hole : SDH) et des trous à fond plat (traduction de flat bottom hole : FBH). L'objectif derrière l'utilisation de plusieurs types de défaut est d'améliorer la capacité des réseaux à pouvoir généraliser le contexte de mesure.

Pour créer les données, des acquisitions type FMC et de type onde plane à zéro-degré ont été générées pour différents types de défauts, tailles et positions dans des pièces de différentes

épaisseurs. Le Tableau 2.1 résume les différentes plages de données utilisées pour les différents paramètres.

Tableau 2.1 Plage des données simulées

Dimensions	Épaisseur (T)	Largeur (W)	Diamètre (D)	Hauteur (H)
Plages de données	40 – 80 mm	32 mm	0.5 -1.5 mm	Incréments de 5 mm

Les schémas de la Figure 2.2 présentent les différents paramètres pour les deux types de défauts utilisés.

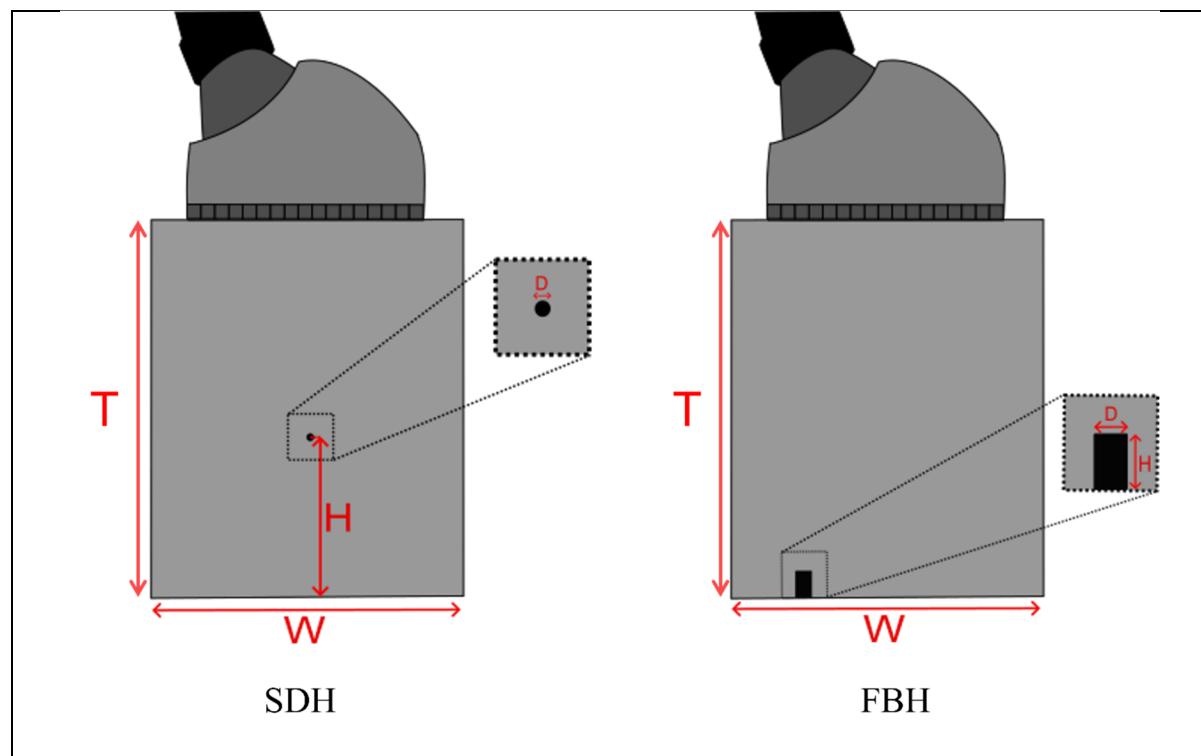


Figure 2.2 Schémas des paramètres utilisés pour les deux types de défauts

Lors de la conception de la base de données, le choix de n'ajouter qu'un seul défaut a été réalisé. Dans les faits, la présence de défauts est assez rare et l'éventualité d'observer une multitude de défauts au même endroit est hautement improbable.

Plus particulièrement, les données simulées ont été créées dans Pogo à partir de modèles 2D discrétisés avec 17 éléments par longueur d'onde. Ensuite, les paramètres définis dans le Tableau 2.1, ont été choisis aléatoirement pour l'ensemble des données simulées.

Ces images ont été modélisées dans un bloc d'acier inoxydable homogène et isotrope aux propriétés définies dans le Tableau 2.2.

Tableau 2.2 Propriétés du matériau simulé

Propriétés du matériau utilisé	Module d'Young	Coefficient de Poisson	Masse volumique
Valeurs	206 GPa	0.29	7850 kg/m ³

Enfin, une sonde multiéléments aux propriétés définies dans le Tableau 2.3 a été modélisée en contact avec la pièce inspectée.

Tableau 2.3 Propriété de la sonde utilisées

Propriétés de la sonde	Fréquence	Bandé passante	Nombre d'éléments	Ouverture	Élévation
Valeurs	5 MHz	75%	64	32 mm	10 mm

Comme énoncé dans le chapitre précédent, afin d'imiter l'allure des données expérimentales, un bruit filtré dans la bande passante de la sonde a été ajouté aux FMC et aux ondes planes à zéro-degré. L'amplitude de ce bruit a été choisi afin de conserver un SNR (ratio signal sur bruit) équivalent aux données expérimentales.

Concernant les données expérimentales, des acquisitions ont été réalisées à l'aide d'un système Verasonics™ Vantage 64 LE et d'une sonde multiéléments Olympus™ de référence 5L64-32X10-A32-P-2.5-OM. La sonde possédant des propriétés identiques à celles énoncées dans le Tableau 2.3 a été couplée aux pièces testées à l'aide d'un gel permettant la transmission des ondes ultrasonores. Enfin, trois blocs différents ont été utilisés afin d'acquérir des données expérimentales ; ces données sont présentées dans le Tableau 2.4.

Tableau 2.4 Plage des données expérimentales

Types de défauts	Matériaux	Épaisseurs (T)	Largeur d'inspection (W)	Diamètre (D)	Hauteur (H)	Nombre d'acquisitions	SNR
SDH	Acier inoxydable	60 mm	32 mm	0.5, 0.8, 1.2, 1.5 mm	Incréments 5 mm	201	~170
	Aluminium	60 mm	32 mm	1 mm	28.6 -30.6 mm	34	~250
FBH	Alliage de Nickel	40, 54, 67, 80 mm	32 mm	0.78 mm	2.54, 3.81 mm	32	~30

Remarque : La position latérale du défaut, pour les données expérimentales et simulées, a été modifiée en déplaçant la sonde avec des incrément de 2mm. Par conséquent, un même défaut a été utilisé pour créer plusieurs acquisitions.

La quantité de données complète utilisée est référencée dans le Tableau 2.5 .

Tableau 2.5 Répartition des données dans les bases de données

Base de données (BDD)	Objectifs	FBH		SDH		
		Alliage de Nickel		Acier inoxydable		Aluminium
		Données simulées	Données expérimentales	Données simulées	Données expérimentales	Données expérimentales
BDD #1	Entraînement et tests	1400	32	1400	201	0
BDD #2	Tests exclusivement	0	0	0	0	34

La base de données (BDD) #1 a été divisée trois parties :

- 80% pour l'entraînement
- 15% pour la validation
- 5% pour le test

La BDD #2 a été exclusivement utilisée pour réaliser des tests de robustesse sur les réseaux de neurones (conditions d'acquisition différentes par rapport à l'entraînement).

Remarque : Il est important de noter que la base d'entraînement est composée à 83% de données simulées alors que les réseaux seront testés sur des données exclusivement expérimentales.

2.2.2 Réseaux de neurones et architectures

Les réseaux utilisés dans cette étude sont des réseaux antagonistes génératifs conditionnels (traduit de conditional Generative Adversarial Networks : cGAN). Leur objectif est d'apprendre une distribution non-linéaire entre une image d'entrée x et une image de sortie y (Isola et al., 2018). Comme énoncé dans la partie précédente (cf. p42), ce réseau est en fait composé de deux sous-réseaux : un modèle Générateur G et un modèle Discriminateur D. Les parties suivantes présentent plus précisément les architectures utilisées pour ces deux modèles.

2.2.2.1 Modèle discriminateur

L'architecture du discriminateur est construite de manière à distinguer les images vraies de la base de données, des images fausses créées par le modèle générateur en se basant sur des caractéristiques présentes dans ces dernières. Pour cette étude, un choix a été réalisé d'utiliser un discriminateur à patch plutôt qu'un discriminateur conventionnel renvoyant une seule réponse pour toute l'image. L'objectif d'un tel modèle est d'obtenir une matrice de classification de taille $N \times N$ correspondant à N^2 régions de pixels (patchs) présentes dans l'image. L'intérêt étant de diminuer le nombre de paramètres du modèle tout en améliorant sa vitesse de traitement et la qualité des résultats renvoyés. L'architecture du discriminateur est composée de couches de convolution, de normalisation de lots (batchs) et de fonctions d'activation type unité linéaire rectifiée avec fuite (leaky rectified linear unit - leaky ReLU) comme présenté sur la Figure 2.3.

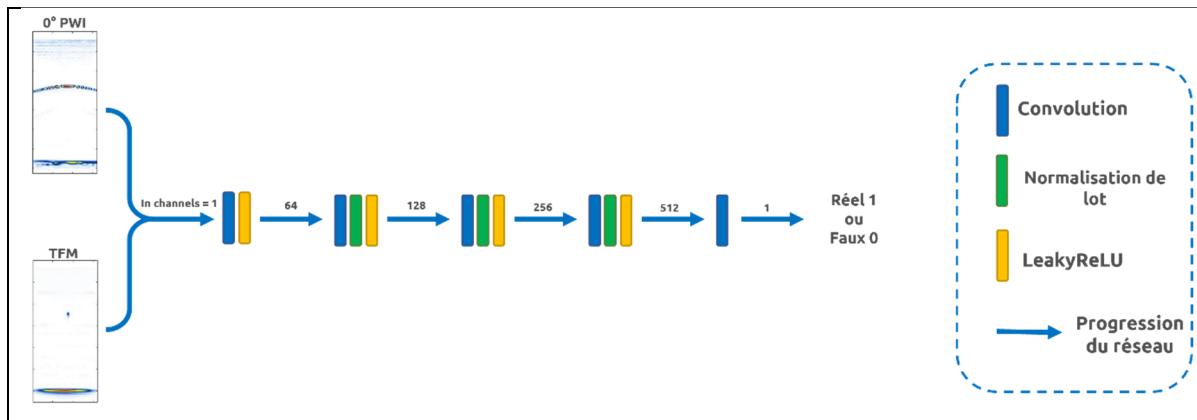


Figure 2.3 Modèle discriminateur

2.2.2.2 Modèles générateurs

Dans cette partie, deux modèles provenant de la littérature seront abordés et comparés. Le générateur est construit de manière à reproduire les images provenant de la base de données, son architecture est basée sur celle de U-Net (Ronneberger et al., 2015), un modèle fortement utilisé dans le cadre de la segmentation des images biomédicales. Son architecture est tout

d'abord composée d'une phase d'extraction des caractéristiques (downsampling phase) puis d'une phase de concaténation des caractéristiques (upsampling phase). Ces deux phases étant composées de couches de convolution, de normalisation de lots (batchs) et de fonctions d'activation type ReLU ou leaky ReLU. De plus, pour éviter le surentraînement, des couches de suppression de neurones ont été ajoutées (dropout) et une fonction d'activation type tangente hyperbolique a été utilisée. Ensuite, afin d'améliorer les performances du réseau, certaines caractéristiques de bas niveau sont directement concaténées aux couches ascendantes comme le présente la Figure 2.4. Ce premier modèle de générateur a été utilisé par Wang et al., (2020).

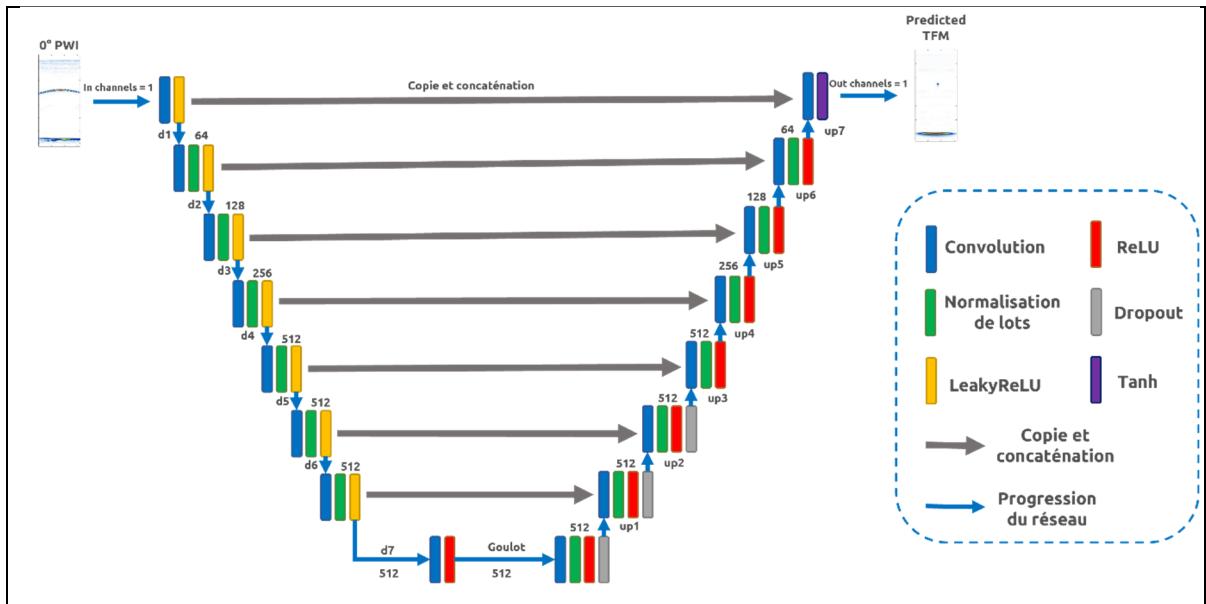


Figure 2.4 Modèle générateur 1

Un second modèle de générateur basé sur les travaux de Tang et al., (2021) a été utilisé dans cette étude. Ce dernier incorpore des couches d'attention supplémentaires dans son architecture afin d'améliorer ses capacités de reconstruction.

Le premier mécanisme d'attention utilisé est l'attention locale (traduit de Local attention) dont l'objectif est d'améliorer la reconstruction des caractéristiques de haute fréquence (Guo et al., 2019). Cette couche peut être comparée à l'ajout d'un masque de netteté : opération classique

utilisée fortement dans la vision par ordinateur, ce masque consiste à ajouter la différence entre l'image originale et une version floutée de cette dernière à l'image originale, ce qui correspond à approximer le résultat renvoyé par un filtre de Laplace (Burt & Adelson, 1987). Cette couche est définie par les opérations successives suivantes :

$$Y_{Blur} = UpSampling(AvgPool(Y_{IN}, k, s, p)) \quad (2.1)$$

avec Y_{IN} le tenseur d'entrée de dimensions $H \times W \times C$ représentant la hauteur, la largeur et le nombre de canaux de l'image (RVB : rouge, vert, bleu) et k , s et p représentant la taille du noyau (traduction de kernel), la taille du pas (stride size) et le taux de remplissage (padding size). Le résultat de cette opération est une version floutée de l'image d'entrée. Une opération de ReLU est ensuite appliquée à la différence avec l'image originale Y_{IN} de façon à approximer le résultat par un filtre de Laplace ou un filtre de différences gaussiennes (Lowe, 2004),

$$Y_R = ReLU(Y_{Blur} - Y_{IN}) \quad (2.2)$$

Enfin, une multiplication de Hadamard est réalisée,

$$Y_{OUT} = Y_{IN}(1 + \beta \times Y_R) \quad (2.3)$$

avec β un paramètre permettant de quantifier l'influence de ce module d'attention.

Le deuxième mécanisme d'attention utilisé est l'attention par pixels (traduit de pixel attention) comme présenté par Zhao et al., (2020). L'objectif de ce module est d'améliorer les caractéristiques de l'image en ajoutant des informations spatiales et d'autres liées aux canaux de l'image. Son expression peut être décrite comme suivant,

$$(f_{1 \times 1}(Y_{k-1})) \otimes Y_{k-1} \quad (2.4)$$

avec $f_{1 \times 1}$ une opération de convolution avec un noyau (kernel) de taille 1×1 , σ une fonction d'activation de type sigmoïde et \otimes le produit de Hadamard (produit termes à termes). Y_k et Y_{k-1} représentent respectivement l'image d'entrée et de sortie de ce module. Ces deux modules d'attention sont ajoutés ensuite à l'architecture précédente comme le présente la Figure 2.5.

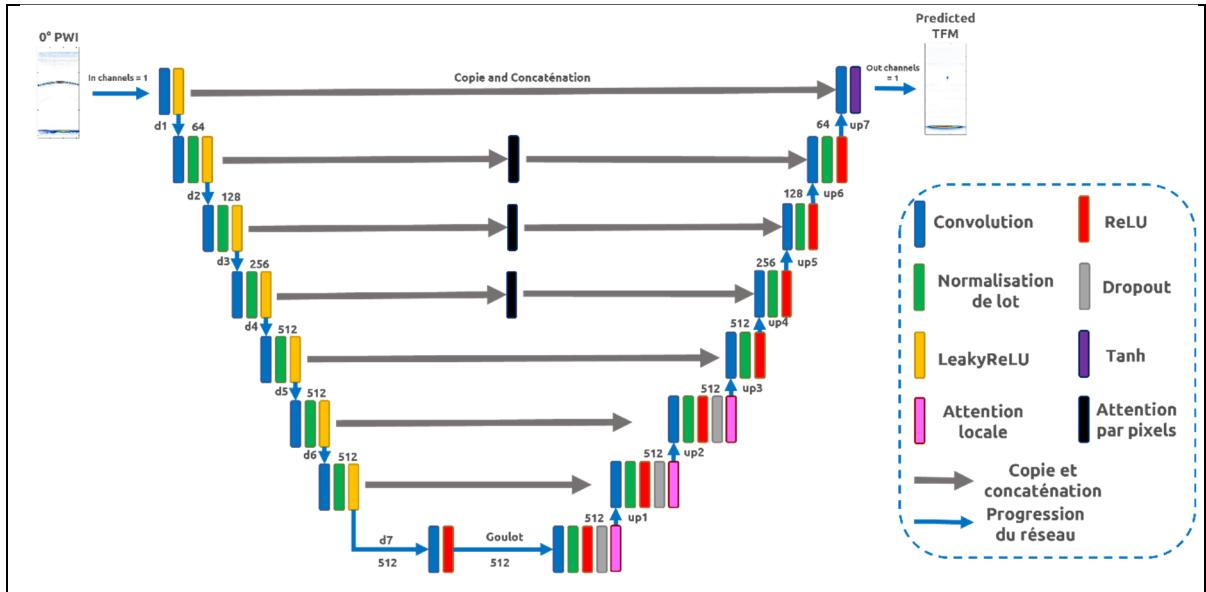


Figure 2.5 Modèle générateur 2

2.2.3 Fonctions de coût

Pour cette étude, une variété de fonctions de coût a été utilisée afin de déterminer leur influence sur l'entraînement du cGAN.

- 1) La fonction de coût adversaire (traduit de adversarial loss) correspondant à l'objectif du cGAN :

$$L_{adv} = E_{x,y}[\log D(x, y)] + E_{x,y}[\log(1 - D(x, G(x, z)))] \quad (2.5)$$

Cette fonction de coût a été définie précédemment dans l'équation 1.25. Pour réaliser cet objectif, une fonction d'entropie croisée binaire a été utilisée comme présenté dans les travaux de Isola et al., (2018).

2) La fonction de coût par pixels (traduit de pixel-wise loss) correspondant à l'écart entre chacun des pixels entre deux images. Cet écart peut être calculé de plusieurs manières, mais les méthodes les plus utilisées restent l'erreur moyenne absolue L_1 (traduit de mean absolute error : MAE) et l'erreur moyenne au carré L_2 (traduit de mean square error : MSE) définies ci-dessous :

$$L_1 = \frac{1}{W \times H \times C} \sum_{i,j,k} |I_{ijk} - I'_{ijk}| \quad (2.6)$$

$$L_2 = \frac{1}{W \times H \times C} \sum_{i,j,k} (I_{ijk} - I'_{ijk})^2 \quad (2.7)$$

avec $W \times H \times C$ représentant la largeur, la hauteur et le nombre de canaux de l'image et I_{ijk} et I'_{ijk} l'intensité des pixels pour une même position (i, j, k) . La fonction de coût L_1 a été utilisée plutôt que L_2 car elle introduit moins de flous dans l'image (Isola et al., 2018).

3) La fonction de coût perceptuelle (traduit de perceptual loss) a été utilisée afin de quantifier les écarts de caractéristiques plutôt que les écarts entre pixels (Johnson et al., 2016). Cette fonction de coût est basée sur l'utilisation du modèle Visual Geometry Group 16 (VGG16) pré-entraîné sur une grande quantité de données génériques à des fins de classification. De ce modèle sont extraites les cartes de caractéristiques (filtres) permettant l'extraction des caractéristiques propres aux images comparées. En considérant $\Phi_j(I)$ la j ème couche d'activation de VGG16 produite par une image I , de dimensions $H_j \times W_j \times C_j$, la fonction de coût perceptuelle se calcule comme suivant,

$$L_{perc}^{\Phi_j}(G(x), y) = \frac{1}{W_j \times H_j \times C_j} \|\Phi_j(y) - \Phi_j(G(x))\|_2^2 \quad (2.8)$$

avec x représentant l'image d'entrée, y l'image de référence et $G(x)$ l'image générée par le réseau. La fonction de coût globale du générateur se formule alors :

$$L_G = \omega L_{adv} + \lambda L_1 + \gamma L_{perc} \quad (2.7)$$

avec ω , λ et γ des coefficients permettant de quantifier l'influence de chacune des fonctions de coût lors de l'entraînement.

2.2.4 Processus d'entraînement

Le processus d'entraînement du cGAN est représenté sur la Figure 2.6,

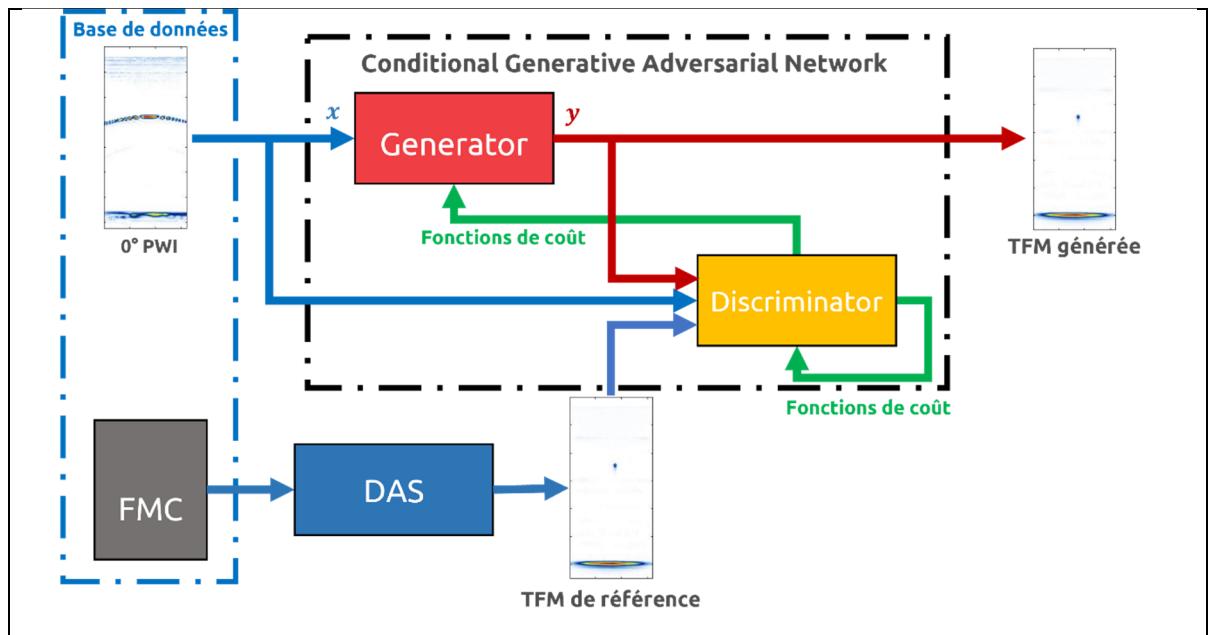


Figure 2.6 Processus d'entraînement

Lors de l'entraînement, pour le calcul des gradients, un Adam Solver (Kingma & Ba, 2017) a été utilisé avec un taux d'apprentissage fixe de 0.001 et des paramètres $\beta_1 = 0.5$, $\beta_2 = 0.999$.

La mise à jour des gradients se faisait par lot de 16 images (batch size) et l'entraînement s'est déroulé sur 250 époques. Le paramètre β dans le mécanisme d'attention a été défini à 0.07.

Concernant les données, avant d'être utilisées dans l'entraînement, ces dernières ont été normalisées par rapport à leur amplitude maximum et redimensionnées en des images de 256×256 .

La création des modèles et script a été effectuée à l'aide du module python Pytorch et l'entraînement a été exécuté sur un serveur linux équipé de deux cartes graphiques NVIDIA™ Quadro P6000. Le Tableau 2.6 présente les différents modèles de cGAN utilisés dans cette étude, basé sur des combinaisons entre architecture du générateur et paramètres des fonctions de coût utilisées.

Tableau 2.6 Modèles de cGAN utilisés

cGAN Modèle	Architecture du générateur	Paramètres		
		ω	λ	γ
Modèle 1	Générateur 1	1	300	0
Modèle 2	Générateur 1	1	3^{-3}	3^{-2}
Modèle 3	Générateur 2	1	3^{-3}	3.8^{-2}

2.2.5 Métriques

Dans cette partie seront présentés les différentes métriques mises en place afin de comparer les modèles de cGAN présentés ci-dessus à la méthode d'imagerie de référence TFM. L'objectif étant de comparer de manière quantitative les images reconstruites.

1) La première métrique définie est l'erreur de position E_{pos} . Cette erreur consiste à calculer la distance Euclidienne entre la position du défaut dans l'image de référence et la position du défaut dans l'image générée afin de s'assurer que ce dernier soit reconstruit au bon endroit. Cette métrique se formule comme suivant,

$$E_{pos} = \frac{1}{N} \sum_i^N \sqrt{(x_i - x'_i)^2 + (y_i - y'_i)^2} \quad (2.8)$$

avec N le nombre de données à tester, (x, y) et (x', y') les coordonnées du défaut respectivement dans la TFM de référence et la TFM générée par le cGAN. E_{pos} correspond à la moyenne des positions de l'ensemble des données testées.

2) La deuxième métrique utilisée est le contraste de signal sur bruit (traduit de contrast to noise ratio : CNR). L'objectif de cette dernière est de déterminer le contraste entre les pixels de la région d'intérêt (le défaut) et le bruit de fond de l'image, cette métrique est décrite comme suivant,

$$CNR = \frac{|\mu_{df} - m_{bg}|}{\sigma_{bg}} \quad (2.9)$$

avec μ_{df} la moyenne des valeurs des pixels de la zone d'intérêt définie à -6dB du maximum d'amplitude, m_{bg} la médiane des valeurs des pixels dans le fond de l'image sans le défaut ni le fond de la pièce et σ_{bg} l'écart type des valeurs dans le fond de l'image sans le défaut ni le fond de la pièce. Ensuite, les valeurs du CNR ont été calculées en décibels (dB),

$$CNR_{dB} = 20 \times \log_{10}(CNR) \quad (2.10)$$

Plus la valeur du CNR est importante, plus le défaut est contrasté par rapport au fond de la pièce. En imagerie ultrasonore, l'objectif est de maximiser ce paramètre. Ensuite, afin de tenir compte des variations d'amplitude du défaut générées par certaines variations dans les

conditions expérimentales ou la position du défaut sous la sonde, la différence de CNR avec l'image de référence a été utilisée,

$$CNR_{diff\ dB} = CNR_{cGAN\ dB} - CNR_{référence\ dB} \quad (2.11)$$

L'objectif est ici de venir quantifier l'écart en décibel entre les deux reconstructions.

3) La troisième métrique utilisée est l'indicateur de performance de la matrice (traduit de Array Performance Indicator API). Cette métrique permet de quantifier l'aire du défaut, elle se calcule comme suivant,

$$API = \frac{A_{-6dB}}{\lambda^2} \quad (2.12)$$

avec λ la longueur d'onde de l'onde ultrasonore propagée dans le matériau et A_{-6dB} la surface totale générée par les pixels d'amplitude supérieure à -6 dB du maximum dans la région d'intérêt : ici le défaut. Pour comparer plus précisément cette différence de reconstruction entre l'image générée par le cGAN et l'image de référence, la différence a été utilisée dans la suite,

$$API_{diff} = API_{cGAN} - API_{référence} \quad (2.13)$$

4) Concernant les vitesses des méthodes, les temps d'acquisition et de reconstruction ont été considérés dans la comparaison. Le temps d'entraînement du réseau n'a quant à lui pas été considéré car ce dernier ne survient qu'en dehors des périodes d'inspection.

5) La taille des données nécessaires aux deux méthodes en entrée a également été considérée.

2.3 Conclusion

Dans cette partie, les objectifs de l'étude et la méthode mise en place a été détaillée. Les bases de données ont été introduites ainsi que les méthodes utilisées pour les obtenir. Les différentes architectures, de cGAN, provenant de la littérature ont été introduites ainsi que les fonctions de coût choisies. Les conditions d'entraînement ont ensuite été décrites et les différents modèles de cGAN ont été présentés. Enfin, les métriques utilisées afin de comparer les différentes méthodes ont été présentées. Le dernier chapitre présente les résultats obtenus.

CHAPITRE 3

RESULTATS ET VALIDATION EXPERIMENTALE

Dans le chapitre précédent, les objectifs et la méthodologie utilisée ont été présentés. Trois modèles de cGAN, introduits dans le Tableau 2.6, vont être comparés dans deux scénarios distincts afin d'en comparer les performances : (1) les données de test sont dans la plage de l'entraînement BDD #1 (2) les données sont en dehors de la plage de l'entraînement BDD #2 comme présenté dans le Tableau 2.5. Dans ce chapitre seront présentés les résultats, puis une discussion sera réalisée avant de conclure quant à la pertinence de la méthode proposée.

Remarque : les données testées de la base de données BDD #1 sont différentes de celles utilisées pour l'entraînement.

3.1 Résultats

Dans cette section seront présentées les reconstructions pour les différents modèles puis des graphes et des tableaux seront utilisés pour comparer les différentes métriques mises en place.

3.1.1 Reconstruction d'images

Dans cette partie, les échelles de couleur des images ont été compressées afin d'augmenter la visibilité des détails présents dans l'image : comme les artéfacts, le bruit de fond de l'image...

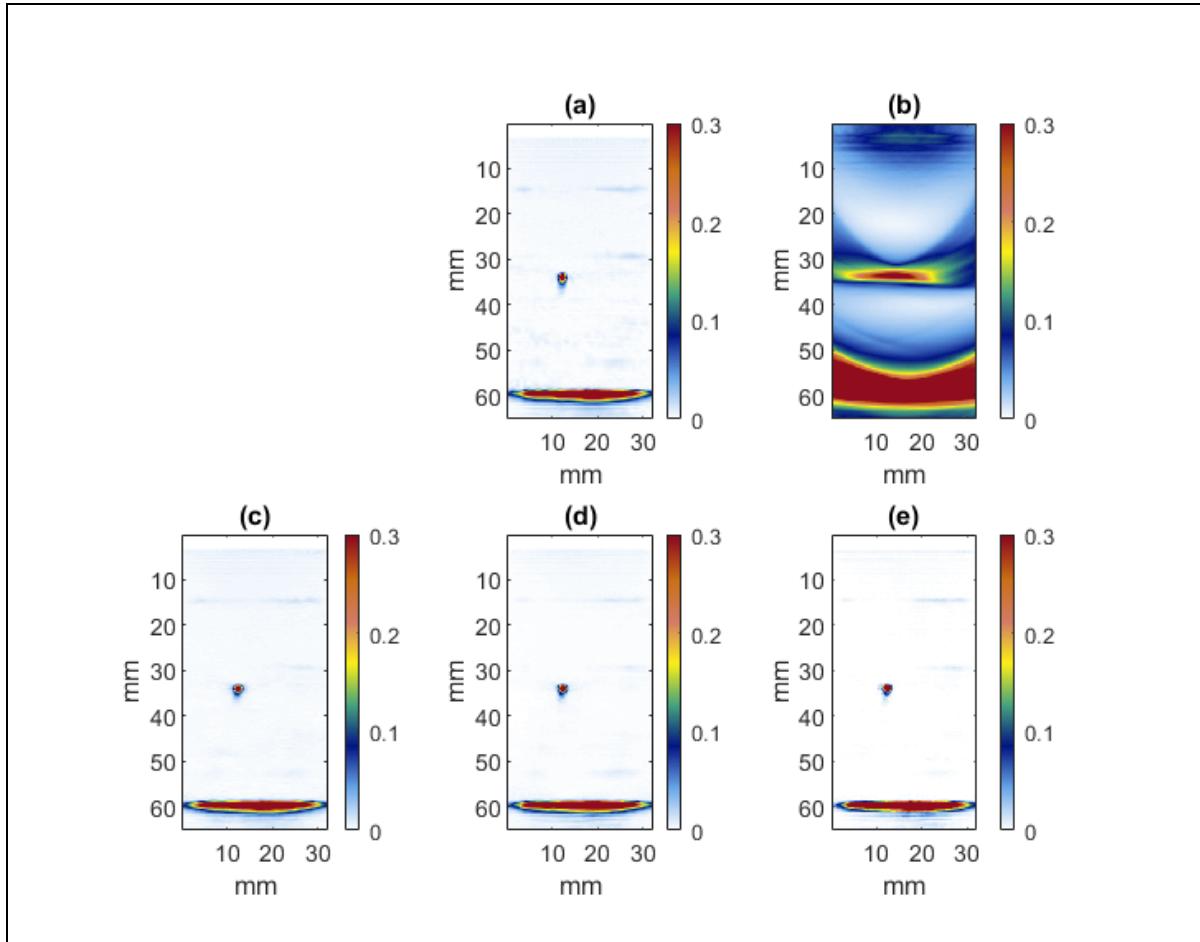


Figure 3.1 Reconstruction d'un même SDH
L'onde plane d'entrée est dans la plage d'entraînement des réseaux

La Figure 3.1 représente différentes reconstructions d'un même défaut type SDH,

- (a) représente l'image de référence obtenue grâce au DAS sur une FMC,
- (b) représente l'image obtenue grâce au DAS sur une onde plane à zéro-degré utilisée comme entrée par le cGAN,
- (c) représente l'image générée par le modèle 1,
- (d) représente l'image générée par le modèle 2,
- (e) représente l'image générée par le modèle 3.

L'onde plane utilisée est dans la plage d'entraînement des cGAN.

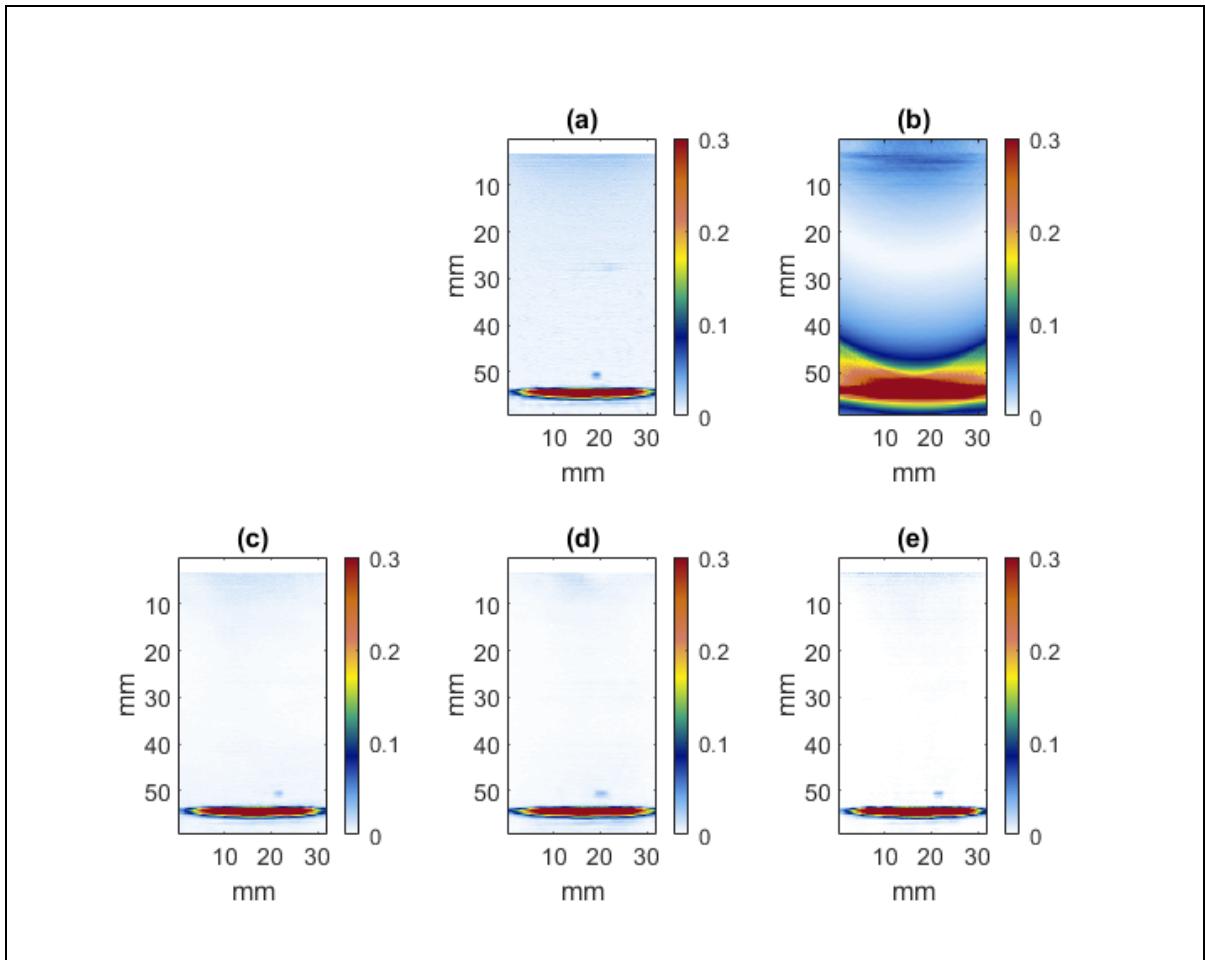


Figure 3.2 Reconstruction d'un même FBH
L'onde plane d'entrée est dans la plage d'entraînement des réseaux

La Figure 3.2 représente différentes reconstructions d'un même défaut type FBH,

- (a) représente l'image de référence obtenue grâce au DAS sur une FMC,
- (b) représente l'image obtenue grâce au DAS sur une onde plane à zéro-degré utilisée comme entrée par le cGAN,
- (c) représente l'image générée par le modèle 1,
- (d) représente l'image générée par le modèle 2,
- (e) représente l'image générée par le modèle 3.

L'onde plane utilisée est dans la plage d'entraînement des cGAN.

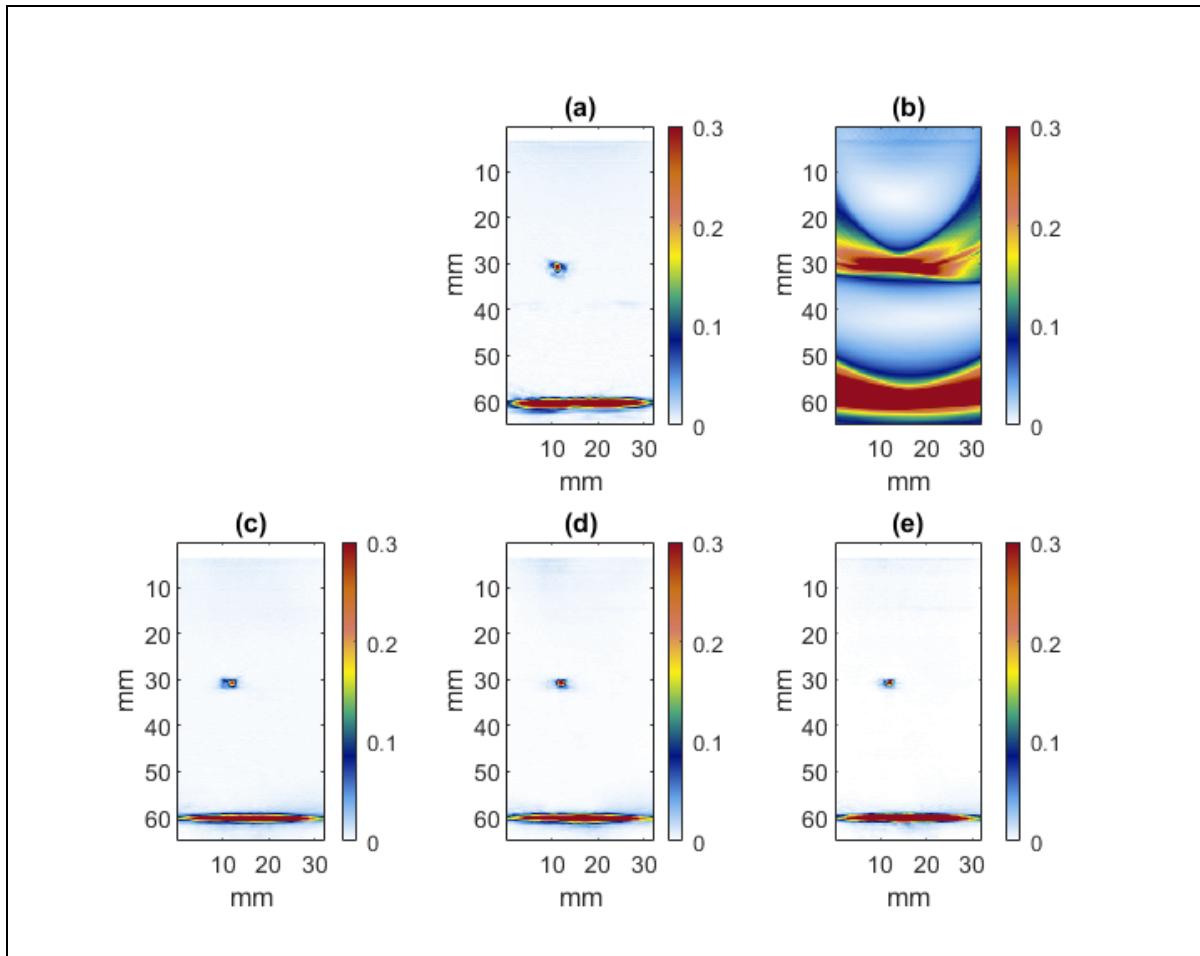


Figure 3.3 Reconstruction d'un même SDH
 L'onde plane d'entrée est en dehors de la plage d'entraînement des réseaux
 le matériau est différent

La Figure 3.3 représente différentes reconstructions d'un même défaut type SDH,

- (a) représente l'image de référence obtenue grâce au DAS sur une FMC,
- (b) représente l'image obtenue grâce au DAS sur une onde plane à zéro-degré utilisée comme entrée par le cGAN,
- (c) représente l'image générée par le modèle 1,
- (d) représente l'image générée par le modèle 2,
- (e) représente l'image générée par le modèle 3.

L'onde plane utilisée est en dehors de la plage d'entraînement des cGAN, le matériau est différent : ici de l'aluminium (cf. Tableau 2.5).

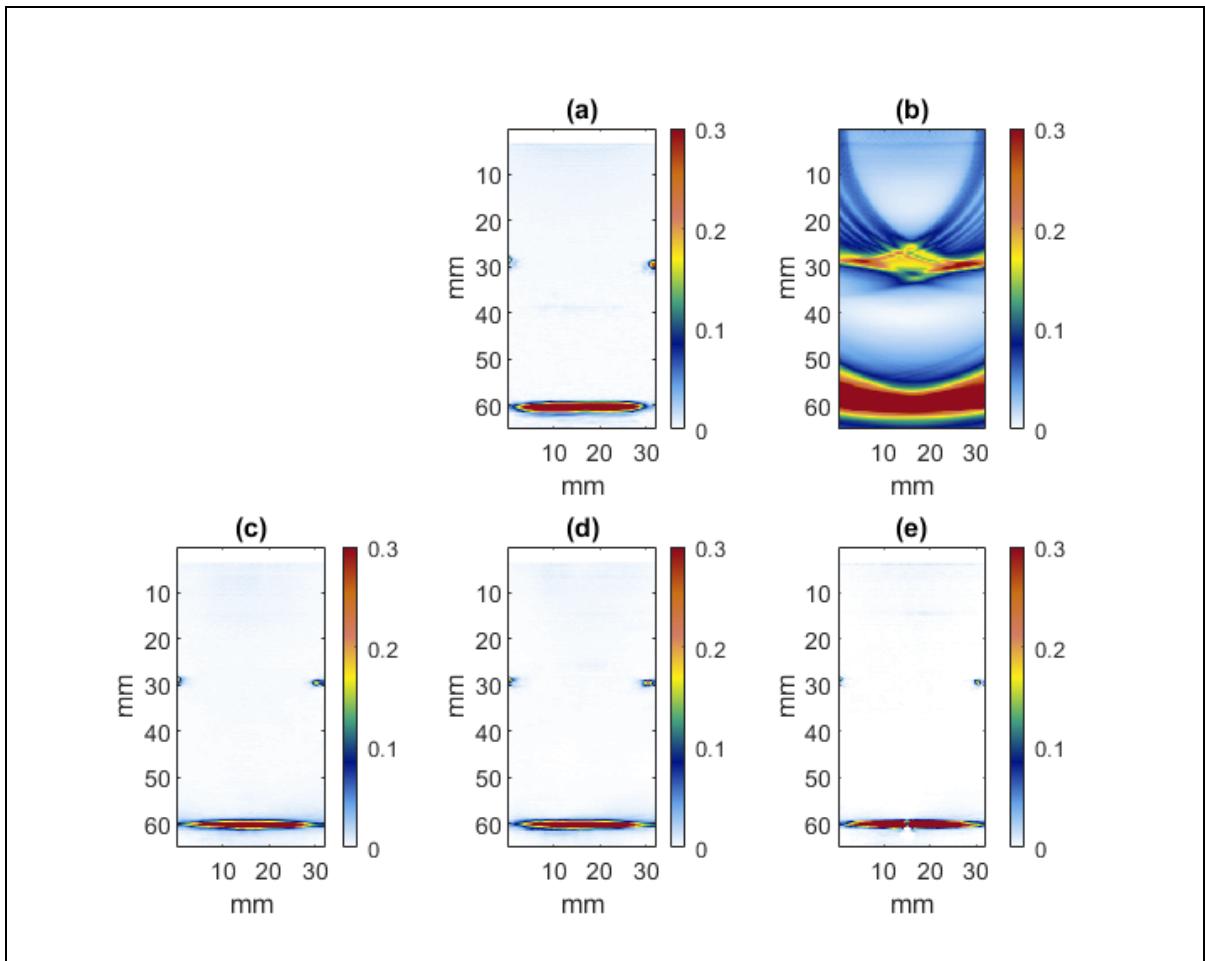


Figure 3.4 Reconstruction de deux SDH

L'onde plane d'entrée est en dehors de la plage d'entraînement des réseaux
le matériau est différent et le nombre de défaut est différent

La Figure 3.4 représente différentes reconstructions de deux défauts type SDH,

- (a) représente l'image de référence obtenue grâce au DAS sur une FMC,
- (b) représente l'image obtenue grâce au DAS sur une onde plane à zéro-degré utilisée comme entrée par le cGAN,
- (c) représente l'image générée par le modèle 1,
- (d) représente l'image générée par le modèle 2,

- (e) représente l'image générée par le modèle 3.

L'onde plane utilisée est en dehors de la plage d'entraînement des cGAN, le matériau est différent : ici de l'aluminium (cf. Tableau 2.5) et le nombre de défauts est différent : l'entraînement était limité à des exemples avec un seul défaut.

3.1.2 Contraste de signal sur bruit (CNR), erreur de position et API

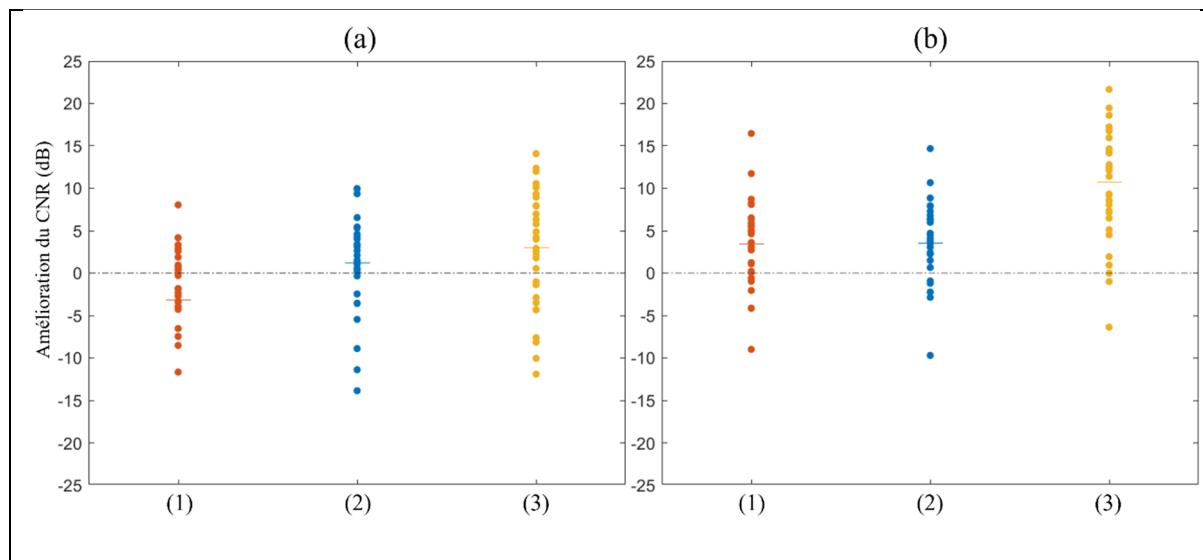


Figure 3.5 Différence de CNR entre la TFM générée et la TFM de référence pour les données testées (a) dans la plage de l'entraînement et (b) en dehors de la plage de l'entraînement

La Figure 3.5 présente la distribution des différences de CNR (cf. Eq 2.11) pour les données testées (a) dans la plage de l'entraînement des cGAN et (b) en dehors de la plage d'entraînement.

- (1) correspond à la distribution du modèle 1,
- (2) correspond à la distribution du modèle 2,
- (3) correspond à la distribution du modèle 3.

Chaque point correspond à une image différente et la ligne horizontale à la moyenne de l'ensemble des données testées.

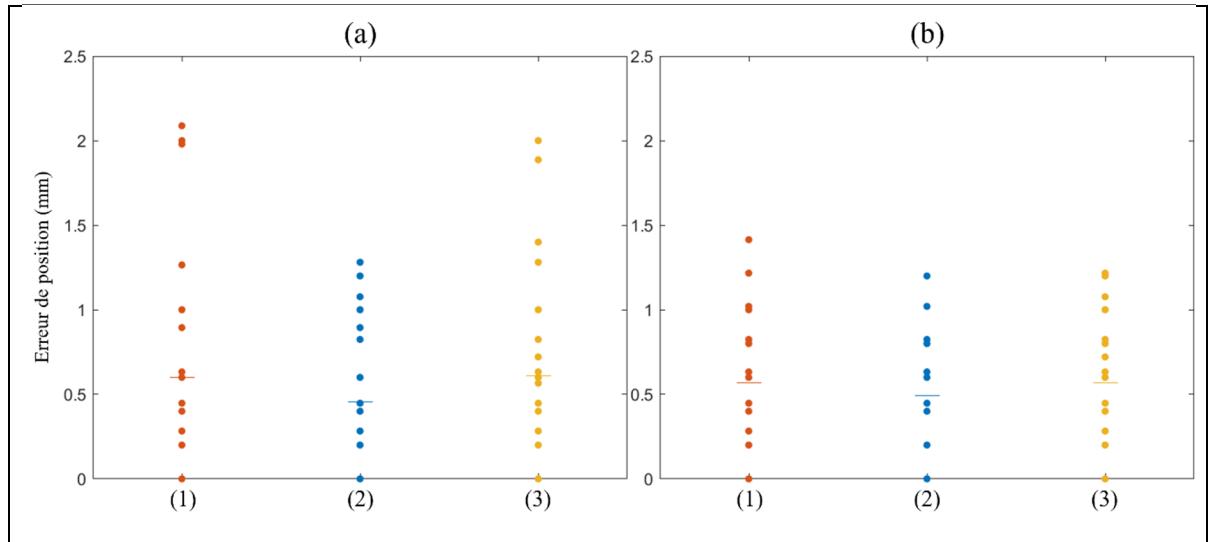


Figure 3.6 Erreur de position du défaut dans la TFM image générée pour les données testées
 (a) dans la plage de l'entraînement et (b) en dehors de la plage de l'entraînement
 La position du défaut dans la TFM de référence est considérée exacte

La Figure 3.6 présente la distribution des erreurs de position des défauts (cf. Eq 2.8) pour les données testées (a) dans la plage de l'entraînement des cGAN et (b) en dehors de la plage d'entraînement.

- (1) correspond à la distribution du modèle 1,
- (2) correspond à la distribution du modèle 2,
- (3) correspond à la distribution du modèle 3.

Chaque point correspond à une image différente et la ligne horizontale à la moyenne de l'ensemble des données testées.

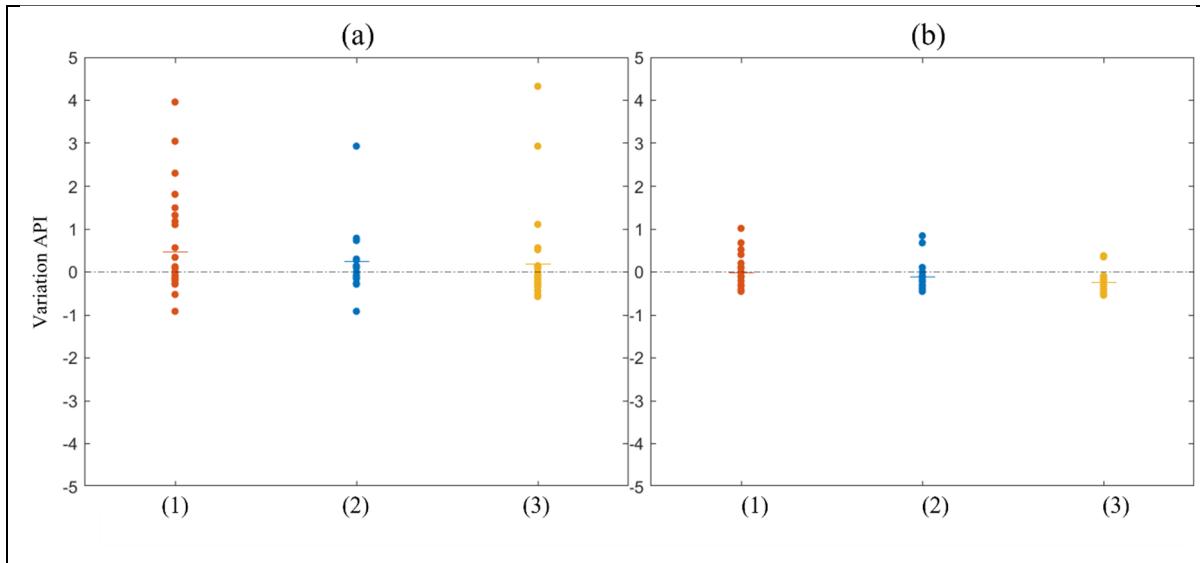


Figure 3.7 Différence d'API entre la TFM image générée et la TFM de référence pour les données testées (a) dans la plage de l'entraînement et (b) en dehors de la plage de l'entraînement

La Figure 3.7 présente la différence d'API entre la TFM générée et la TFM de référence (cf. Eq 2.13) pour les données testées (a) dans la plage de l'entraînement des cGAN et (b) en dehors de la plage d'entraînement.

- (1) correspond à la distribution du modèle 1,
- (2) correspond à la distribution du modèle 2,
- (3) correspond à la distribution du modèle 3.

Chaque point correspond à une image différente et la ligne horizontale à la moyenne de l'ensemble des données testées.

Les valeurs des métriques sont référencées dans le Tableau 3.1 pour les deux différents scénarios de test.

Tableau 3.1 Valeurs caractéristiques pour la différence de CNR, l'erreur de position et l'API

Models	Tested data within the training range						Tested data outside of the training range					
	CNR difference (dB)		Position error E_{pos} (mm)		API difference		CNR difference (dB)		Position error E_{pos} (mm)		API difference	
	mean	STD	mean	STD	mean	STD	mean	STD	mean	STD	mean	STD
cGAN model 1	-3.17	8.27	0.38	0.29	0.46	1.1	3.43	5.0	0.48	0.33	-0.019	0.36
cGAN model 2	1.26	5.61	0.34	0.22	0.25	1.1	3.52	4.74	0.49	0.28	-0.12	0.31
cGAN model 3	3.02	6.92	0.44	0.34	0.20	1	10.72	7.14	0.55	0.35	-0.25	0.21

3.1.3 Vitesse des modèles et taille des données

Le Tableau 3.2 présente une comparaison entre le modèle proposé incorporant un cGAN et la méthode conventionnelle utilisant le DAS et une acquisition type FMC en termes de temps de traitement et de taille de données en entrée.

Tableau 3.2 Comparaison des performances entre le DAS et la méthode alternative utilisant un cGAN

Méthode	Entrée	Temps d'acquisition (entrée)	Taille de la donnée d'entrée	Temps de reconstruction (TFM)	Temps total
cGAN	0° PW	25 μ s	160 kB	0.05 s	~ 50 ms
DAS	FMC	1.6 ms	12 MB	6 s	~ 6 s

3.2 Discussion préliminaire

Pour concevoir une méthode par cGAN efficace et capable de reconstruire des TFM équivalentes à celles obtenues par le DAS avec des acquisitions type FMC, il a été important de s'assurer qu'il existait bien un lien physique entre une acquisition type onde plane et une

TFM reconstruite. En effet, avant toutes utilisations de réseaux de neurones il intéressant de voir si la tâche est au préalable réalisable par exemple par un humain. La Figure 2.1 est intéressante puisqu'elle présente ces deux résultats pour un même SDH. Ici, l'écho représenté sur l'image de l'onde plane correspond à l'écho généré par le défaut, sa position et sa forme dépendent respectivement de la position du défaut et de sa morphologie (taille, type ...). La position du défaut peut être donc estimée grâce au maximum dans cet écho comme on l'observe avec l'image de la TFM. Pour cette raison, on peut conclure sur la faisabilité de l'entraînement d'un point de vue physique, et donc sur la capacité des filtres de convolution à pouvoir extraire ces informations.

3.3 Discussion et validation des résultats

Trois modèles de cGAN ont été créés et comparés dans deux cas d'utilisation. Cette section vise à discuter de ces derniers dans le cadre d'une imagerie à haute cadence et à définir de leurs capacités à pouvoir remplacer ou non les méthodes usuelles comme le DAS.

Tout d'abord, il est important de noter que les modèles de cGAN ont été entraînés sur des données majoritairement simulées. En effet, environ 92% de la base de données était composée de simulations. Ensuite, dans tous les scénarios de cette étude, les images reconstruites par les cGAN avaient une meilleure résolution et un meilleur contraste que celles reconstruites par le DAS pour une même entrée : une onde plane à zéro-degré. Cette constatation est visible dans les Figure 3.1 à Figure 3.4, entre l'image (b) et les images (c), (d) et (e). En effet, dans l'image (b) il est difficile de localiser précisément le défaut et le fond de la pièce, ce qui n'est pas le cas des images générées par les cGAN (c), (d) et (e). Pour cette raison, il n'a pas été jugé nécessaire de considérer l'image (b) avec les autres métriques.

Ensuite, concernant les capacités de reconstruction des modèles de cGAN, les images générées par le modèle 1 (c) sont toujours les plus proches visuellement des images de référence (a) correspondant aux reconstructions par la méthode du DAS sur des FMC, principalement concernant la reconstruction des artefacts et du bruit de fond. Cette capacité de reconstruction

s'explique par l'utilisation de la fonction de coût par pixels qui a une grosse influence sur ce modèle (un coefficient de 300 cf. Tableau 2.6). En effet, l'objectif de cette fonction est de minimiser l'écart entre l'image de référence et l'image générée en considérant de manière équivalente chacun des pixels. De ce fait, du point de vue strict de l'entraînement, ces images doivent être considérées comme des reconstructions de bonne qualité. Cependant, du point de vue du CND, le bruit et les artefacts sont néfastes et peuvent gêner dans la détection de défauts. Par conséquent, afin de s'intéresser aux caractéristiques plus importantes, les modèles de cGAN 2 et 3 ont été créés. Plutôt que de viser une équivalence parfaite entre l'image d'entraînement et l'image générée, l'utilisation de la fonction de coût perceptuelle et des mécanismes d'attention avait pour but d'améliorer le processus de reconstruction dans les zones d'intérêt : défauts et fond de la pièce. Les images (d) et (e) des Figure 3.1 à Figure 3.4 présentent les avantages de cette nouvelle stratégie : la présence d'artefact et de bruit de fond se retrouve fortement atténuée et davantage pour le modèle 3 (e) grâce aux mécanismes d'attention.

Cette différence se retrouve dans la Figure 3.5 avec la représentation des distributions de CNR. En effet, dans le graphe (a), le modèle 1 ne présente pas de réelles améliorations, la moyenne de la distribution est centrée sur un CNR inférieur à celui des images de référence (-3.17 dB). Cet exemple montre que le modèle 1 ne parvient pas à dépasser les performances en termes de CNR du DAS lorsque la principale fonction de coût utilisée ne s'intéresse qu'à la différence entre chaque pixel. À l'inverse pour les modèles 2 et 3, ces distributions sont centrées sur des valeurs supérieures à zéro, ce qui montre une amélioration moyenne des CNR par rapport aux images de référence. La fonction de coût perceptuelle et les mécanismes d'attention ont eu pour effet d'ajouter un biais dans l'entraînement permettant ainsi d'améliorer les capacités de reconstruction des modèles. Ce résultat est particulièrement intéressant si l'on considère le fait qu'il a suffi d'une seule onde plane pour reconstruire ces TFM ayant un meilleur contraste que la TFM générée à l'aide d'une FMC.

Il est ensuite essentiel de noter que l'importante plage de données générées grâce à la création aléatoire des données simulées a permis aux modèles de généraliser le contexte d'étude. En

effet, dans les Figure 3.3 et Figure 3.4, les modèles sont parvenus à recréer des TFM d'aussi bonne qualité que les TFM de référence. Les trois modèles sont parvenus à reconstruire la section d'un bloc d'aluminium non utilisé dans l'entraînement. La longueur d'onde étant différente, les proportions de l'image d'entrée se retrouvaient changées. Cette capacité de généralisation se retrouve également dans la Figure 3.4 où les modèles ont été capables de reconstruire deux SDH alors que l'entraînement n'était limité qu'à des exemples d'acquisition présentant un unique défaut. Les modèles sont donc parvenus à séparer les influences des deux défauts pour les reconstruire fidèlement.

De plus, cette capacité à généraliser des modèles ne s'arrête pas à de simples reconstructions. En effet, les Figure 3.5 et Figure 3.7 montrent que les modèles 1, 2 et 3 fonctionnent encore mieux sur les nouvelles données (en dehors de la plage d'entraînement). Le CNR ainsi que l'API sont fortement améliorés par rapport au test sur des données proches de l'entraînement. Tout d'abord, la moyenne du CNR est augmentée pour les trois modèles et l'écart type diminue pour les modèles 1 et 2. Ensuite, dans le cas de l'API, cette dernière tend à s'améliorer avec l'utilisation des nouvelles données. En effet, dans le cas des données proches de l'entraînement, le cGAN a tendance à surestimer la taille du défaut et donc augmenter les valeurs de la différence d'API, ce qui n'est pas le cas sur les nouvelles données où le cGAN propose une meilleure estimation de la taille du défaut car les valeurs sont centrées sur 0, ce qui correspond à une surface identique du défaut avec l'image de référence. Néanmoins, sur la Figure 3.6, une diminution de la précision dans le positionnement est observable avec une augmentation de l'erreur moyenne de positionnement ainsi que de l'écart-type.

Le modèle 3 quant à lui, profitant des couches d'attention, voit ses performances s'améliorer grandement surtout vis-à-vis du CNR : environ 93% des données ont un meilleur CNR que les images de référence obtenues grâce au DAS sur des FMC, ce gain est en moyenne de 10.7 dB. Une des raisons possibles à de tels progrès de reconstructions est une meilleure qualité des données testées ; en effet, comme présenté dans le Tableau 2.4 les données testées ont un meilleur ratio signal sur bruit (SNR) que les données expérimentales de l'entraînement. Cette raison encourage l'idée que les performances des réseaux sont avant tout corrélées à la qualité

des données proposées. Le modèle semble également avoir développé une meilleure compréhension des caractéristiques importantes de l'image : ce dernier s'intéresse davantage à la reconstruction du défaut et du fond de la pièce plutôt qu'au bruit de fond de l'image et aux artefacts améliorant ainsi le CNR.

Concernant la vitesse des méthodes, le cGAN est complètement supérieur à la méthode utilisant le DAS et l'acquisition type FMC comme le présente le Tableau 3.2. En effet, l'utilisation d'une acquisition type onde plane permet de réduire par 64 le temps d'acquisition et la taille des données utilisées par 75 environ. Le temps de reconstruction a également été diminué par un facteur de 120 environ. À noter qu'il est plus important de diminuer le temps d'acquisition plutôt que le temps de traitement de la méthode, car ce dernier peut toujours être amélioré grâce aux optimisations électroniques, alors que le premier constitue une limite physique (temps de propagation des ultrasons dans les matériaux).

3.4 Conclusion et validation des résultats

La méthode alternative proposée présente de nombreux avantages dans le cadre d'une inspection haute cadence par rapport aux méthodes conventionnels type DAS. En effet, cette méthode propose des images de résolution équivalentes au DAS sans se contraindre à une acquisition type FMC et cela dans des temps drastiquement réduits. Cette méthode alternative est également intéressante grâce à l'utilisation des fonctions de coût qui dans certains cas d'utilisation pourraient permettre d'améliorer les images reconstruites grâce à des biais introduits dans l'entraînement. Néanmoins, il est important de noter que cette méthode alternative nécessite un entraînement et qu'elle ne peut convenir qu'à des cas d'inspection bien maîtrisés.

CONCLUSION

Dans cette étude, une méthode alternative incorporant des réseaux antagonistes génératifs conditionnels (cGAN) a été présentée afin de contourner la méthode conventionnelle de somme et de délais (DAS) fortement utilisée dans le cadre de l'imagerie par ultrasons industrielle.

Pour ce faire, une étude approfondie de la littérature a tout d'abord été réalisée, et les concepts liés à l'imagerie ultrasonore et aux réseaux de neurones artificiels ont été introduit. Ensuite une seconde partie a permis de définir plus concrètement l'objectif de l'étude ainsi que la méthode mise en œuvre pour comparer ces deux méthodes. Enfin, les deux méthodes ont été testées puis confrontées dans plusieurs cas d'utilisation dans le cadre d'une imagerie haute cadence.

La méthode alternative incorporant le cGAN s'est révélée être pertinente et plus adaptée que la méthode conventionnelle utilisant le DAS dans un contexte d'imagerie haute cadence. En effet, dans un contexte maîtrisé où les composants à tester ont des propriétés et des formes similaires, il est possible d'envisager un modèle comme le cGAN pour venir accélérer le processus de contrôle. Cependant, il reste important de s'assurer que les pièces testées ne s'éloignent pas trop des conditions d'entraînement pour ne pas faire face à des résultats complètement erronés.

RECOMMANDATIONS

Dans cette étude, une méthode incorporant un cGAN a été choisie pour remplacer certaines méthodes usuelles dans un contexte d'imagerie haute cadence. Cependant, de nombreux travaux et améliorations restent encore réalisables.

Tout d'abord, comme présenté dans la littérature, de nombreuses architectures de réseaux de neurones existent et pourraient également être utilisées dans ce même contexte. Une comparaison des différents modèles pourrait donc apporter une valeur ajoutée à cette étude.

Ensuite, faute de temps il n'a pas été possible d'optimiser à 100% les modèles de cette étude. De ce fait, il reste encore largement possible d'améliorer les résultats et donc les écarts avec les méthodes d'imagerie conventionnelles. La base de données peut quant à elle être aussi nettement améliorée, notamment grâce à l'utilisation d'autres sondes, de sabots, de défauts plus complexes...

Enfin, le cas d'étude peut également être plus orienté, par exemple pour l'inspection de filières ou de tuyaux créés à haute cadence. Afin d'accéder à des données plus réelles.

BIBLIOGRAPHIE

- Asl, B. M., & Mahloojifar, A. (2009). Minimum variance beamforming combined with adaptive coherence weighting applied to medical ultrasound imaging. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 56(9), 1923-1931. <https://doi.org/10.1109/TUFFC.2009.1268>
- Burgos-Artizzu, X. P., Coronado-Gutiérrez, D., Valenzuela-Alcaraz, B., Bonet-Carne, E., Eixarch, E., Crispi, F., & Gratacós, E. (2020). Evaluation of deep convolutional neural networks for automatic classification of common maternal fetal ultrasound planes. *Scientific Reports*, 10(1), Art. 1. <https://doi.org/10.1038/s41598-020-67076-5>
- Burt, P. J., & Adelson, E. H. (1987). The Laplacian Pyramid as a Compact Image Code. In M. A. Fischler & O. Firschein (Éds.), *Readings in Computer Vision* (p. 671-679). Morgan Kaufmann. <https://doi.org/10.1016/B978-0-08-051581-6.50065-9>
- Cheeke, J. D. N., & Cheeke, J. D. N. (2002). *Fundamentals and Applications of Ultrasonic Waves*. CRC Press. <https://doi.org/10.1201/9781420042139>
- Cook, R. D., Malkus, D. S., Plesha, M. E., & Witt, R. J. (2007). *Concepts and Applications of Finite Element Analysis*. John Wiley & Sons, Inc.
- Couture, O., Fink, M., & Tanter, M. (2012). Ultrasound contrast plane wave imaging. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 59(12), 2676-2683. <https://doi.org/10.1109/TUFFC.2012.2508>
- Drozdz, M., Moreau, L., Castaings, M., Lowe, M. J. S., & Cawley, P. (2006). Efficient Numerical Modelling of Absorbing Regions for Boundaries Of Guided Waves Problems. *IPConferenceProceedings*, 820(1), 126-133. <https://doi.org/10.1063/1.2184520>
- Dupont-Marilla, F. (2018). *Inspection de blocs d'acier forgé de grand volume par ultrasons multi-éléments* [Phd, École de technologie supérieure]. <https://espace.etsmtl.ca/id/eprint/2158/>
- Durand, T. (2017). *Weakly supervised learning for visual recognition* [These de doctorat, Paris 6]. <https://www.theses.fr/2017PA066142>
- Guo, J., Ma, S., & Guo, S. (2019). *MAANet : Multi-view Aware Attention Networks for Image Super-Resolution*(arXiv:1904.06252).arXiv. <https://doi.org/10.48550/arXiv.1904.06252>

- Huthwaite, P. (2014). Accelerated finite element elastodynamic simulations using the GPU. *Journal of Computational Physics*, 257, 687-707. <https://doi.org/10.1016/j.jcp.2013.10.017>
- Isola, P., Zhu, J.-Y., Zhou, T., & Efros, A. A. (2018). Image-to-Image Translation with Conditional Adversarial Networks. *arXiv:1611.07004 [cs]*. <http://arxiv.org/abs/1611.07004>
- Johnson, J., Alahi, A., & Fei-Fei, L. (2016). *Perceptual Losses for Real-Time Style Transfer and Super-Resolution* (arXiv:1603.08155). arXiv. <https://doi.org/10.48550/arXiv.1603.08155>
- Kanopoulos, N., Vasanthavada, N., & Baker, R. L. (1988). Design of an image edge detection filter using the Sobel operator. *IEEE Journal of Solid-State Circuits*, 23(2), 358-367. <https://doi.org/10.1109/4.996>
- Kingma, D. P., & Ba, J. (2017). *Adam: A Method for Stochastic Optimization* (arXiv:1412.6980). arXiv. <https://doi.org/10.48550/arXiv.1412.6980>
- Le, K. (2021, décembre 8). An overview of VGG16 and NiN models. *MLearning.Ai*. [https://medium.com/mlarning-ai/an-overview-of-vgg16-and-nin-models-96e4bf398484](https://medium.com/mlearning-ai/an-overview-of-vgg16-and-nin-models-96e4bf398484)
- Le Manuel numérique max—Belin Education.* (s. d.). Consulté 13 octobre 2022, à l'adresse https://manuelnumeriquemax.belin.education/svt-premiere/topics/svt1-ch07-120-a_seismes-et-ondes-sismiques
- Li, B., & Xu, P. (s. d.). Convolutional neural networks. *Neural Networks*, 26.
- Li, Q., Cai, W., Wang, X., Zhou, Y., Feng, D. D., & Chen, M. (2014). Medical image classification with convolutional neural network. *2014 13th International Conference on Control Automation Robotics & Vision (ICARCV)*, 844-848. <https://doi.org/10.1109/ICARCV.2014.7064414>
- Li, Y.-C., Shen, T.-Y., Chen, C.-C., Chang, W.-T., Lee, P.-Y., & Huang, C.-C. J. (2021). Automatic Detection of Atherosclerotic Plaque and Calcification From Intravascular Ultrasound Images by Using Deep Convolutional Neural Networks. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 68(5), 1762-1772. <https://doi.org/10.1109/TUFFC.2021.3052486>
- Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., van der Laak, J. A. W. M., van Ginneken, B., & Sánchez, C. I. (2017). A survey on deep learning in medical image analysis. *Medical Image Analysis*, 42, 60-88. <https://doi.org/10.1016/j.media.2017.07.005>

- Lowe, D. G. (2004). Distinctive Image Features from Scale-Invariant Keypoints. *International Journal of Computer Vision*, 60(2), 91-110. <https://doi.org/10.1023/B:VISI.0000029664.99615.94>
- Mathew, J. (2022, août 1). *Deep Learning for Image segmentation*. Medium. <https://medium.datadriveninvestor.com/deep-learning-for-image-segmentation-d10d19131113>
- Mikołajczyk, A., & Grochowski, M. (2018). Data augmentation for improving deep learning in image classification problem. *2018 International Interdisciplinary PhD Workshop (IIPhDW)*, 117-122. <https://doi.org/10.1109/IIPHDW.2018.8388338>
- Montaldo, G., Tanter, M., Bercoff, J., Benech, N., & Fink, M. (2009). Coherent plane-wave compounding for very high frame rate ultrasonography and transient elastography. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 56(3), 489-506. <https://doi.org/10.1109/TUFFC.2009.1067>
- Munir, N., Kim, H.-J., Park, J., Song, S.-J., & Kang, S.-S. (2019). Convolutional neural network for ultrasonic weldment flaw classification in noisy conditions. *Ultrasonics*, 94, 74-81. <https://doi.org/10.1016/j.ultras.2018.12.001>
- Nair, A. A., Tran, T. D., Reiter, A., & Lediju Bell, M. A. (2018). A Deep Learning Based Alternative to Beamforming Ultrasound Images. *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 3359-3363. <https://doi.org/10.1109/ICASSP.2018.8461575>
- Pavlov, I., Prado, E., Navab, N., & Zahnd, G. (2019). Towards in-vivo ultrasound-histology : Plane-waves and generative adversarial networks for pixel-wise speed of sound reconstruction. *2019 IEEE International Ultrasonics Symposium (IUS)*, 1913-1916. <https://doi.org/10.1109/ULTSYM.2019.8925722>
- Perdios, D., Vonlanthen, M., Martinez, F., Ardit, M., & Thiran, J.-P. (2022). CNN-Based Image Reconstruction Method for Ultrafast Ultrasound Imaging. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 69(4), 1154-1168. <https://doi.org/10.1109/TUFFC.2021.3131383>
- Perrot, V., Polichetti, M., Varray, F., & Garcia, D. (2021). So you think you can DAS? A viewpoint on delay-and-sum beamforming. *Ultrasonics*, 111, 106309. <https://doi.org/10.1016/j.ultras.2020.106309>
- Qi, Y., Guo, Y., & Wang, Y. (2021). Image Quality Enhancement Using a Deep Neural Network for Plane Wave Medical Ultrasound Imaging. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 68(4), 926-934. <https://doi.org/10.1109/TUFFC.2020.3023154>

Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net : Convolutional Networks for Biomedical Image Segmentation. In N. Navab, J. Hornegger, W. M. Wells, & A. F. Frangi (Éds.), *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015* (p. 234-241). Springer International Publishing. https://doi.org/10.1007/978-3-319-24574-4_28

Sai Sundar, K. V., Rajamani, K. T., & Siva Sankara Sai, S. (2019). Exploring Image Classification of Thyroid Ultrasound Images Using Deep Learning. In D. Pandian, X. Fernando, Z. Baig, & F. Shi (Éds.), *Proceedings of the International Conference on ISMAC in Computational Vision and Bio-Engineering 2018 (ISMAC-CVB)* (p. 1635-1641). Springer International Publishing. https://doi.org/10.1007/978-3-030-00665-5_151

Schmerr, L. W. (2015). *Fundamentals of ultrasonic phased arrays*. Springer.

Shull, P. (2002). *Nondestructive Evaluation : Theory, Techniques, and Applications*. pp 142. CRC Press. <https://doi.org/10.1201/9780203911068>

Synnevag, J., Austeng, A., & Holm, S. (2009). Benefits of minimum-variance beamforming in medical ultrasound imaging. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 56(9), 1868-1879. <https://doi.org/10.1109/TUFFC.2009.1263>

Tang, J., Zou, B., Li, C., Feng, S., & Peng, H. (2021). Plane-Wave Image Reconstruction via Generative Adversarial Network and Attention Mechanism. *IEEE Transactions on Instrumentation and Measurement*, 70, 1-15. <https://doi.org/10.1109/TIM.2021.3087819>

Virupakshappa, K., Marino, M., & Oruklu, E. (2018). A Multi-Resolution Convolutional Neural Network Architecture for Ultrasonic Flaw Detection. *2018 IEEE International Ultrasonics Symposium (IUS)*, 1-4. <https://doi.org/10.1109/ULTSYM.2018.8579888>

Wang, Y., Kempski, K., Kang, J. U., & Bell, M. A. L. (2020). A Conditional Adversarial Network for Single Plane Wave Beamforming. *2020 IEEE International Ultrasonics Symposium (IUS)*, 1-4. <https://doi.org/10.1109/IUS46767.2020.9251729>

Weston, J., Ratle, F., & Collobert, R. (2008). Deep learning via semi-supervised embedding. *Proceedings of the 25th international conference on Machine learning*, 1168-1175. <https://doi.org/10.1145/1390156.1390303>

Yap, M. H., Pons, G., Martí, J., Ganau, S., Sentís, M., Zwiggelaar, R., Davison, A. K., & Martí, R. (2018). Automated Breast Ultrasound Lesions Detection Using Convolutional Neural Networks. *IEEE Journal of Biomedical and Health Informatics*, 22(4), 1218-1226. <https://doi.org/10.1109/JBHI.2017.2731873>

- Zhao, H., Kong, X., He, J., Qiao, Y., & Dong, C. (2020). Efficient Image Super-Resolution Using Pixel Attention. In A. Bartoli & A. Fusiello (Eds.), *Computer Vision – ECCV 2020 Workshops* (p. 56-72). Springer International Publishing. https://doi.org/10.1007/978-3-030-67070-2_3
- Zhou, R., Fenster, A., Xia, Y., Spence, J. D., & Ding, M. (2019). Deep learning-based carotid media-adventitia and lumen-intima boundary segmentation from three-dimensional ultrasound images. *Medical Physics*, 46(7), 3180-3193. <https://doi.org/10.1002/mp.13581>
- Zhuang, Z., Raj, A. N. J., Jain, A., Ruban, N., Chaurasia, S., Li, N., Lakshmanan, M., & Murugappan, M. (2019). Nipple Segmentation and Localization Using Modified U-Net on Breast Ultrasound Images. *Journal of Medical Imaging and Health Informatics*, 9(9), 1827-1837. <https://doi.org/10.1166/jmhi.2019.2828>