

Chapitre



Applications des approches SVM et SVR pour la classification

III.1. Introduction

Les recherches dans le domaine de la reconstruction d'objets à partir de champs électromagnétiques ont pris une ampleur considérable durant la dernière décennie, en raison de leurs diverses applications dans de nombreux domaines comme la médecine (détection du cancer), la biologie, géophysique et d'autres sciences [62].

Le cancer du sein est l'une des maladies les plus dangereuses dans le monde moderne avec une mortalité élevée [63]. On estime que dans l'Union européenne, toutes les 2,5 minutes, une femme reçoit un diagnostic de cancer du sein. Toutes les 7,5 minutes, une femme meurt de la maladie. L'Algérie a connu ces dix dernières années une évolution importante du nombre de cancers du sein, En 2008 on dénombre 5 000 à 7 000 nouveaux cas alors qu'en 2003 on en comptait 2 000 par an [64].

La détection précoce des petites tumeurs du sein est absolument nécessaire pour prévenir les métastases du cancer et pour l'obtention d'un bon pronostic du traitement du cancer. Par conséquent, il est très important de disposer des technologies et des algorithmes qui sont capables de détecter des structures tumorales de quelques millilitres dans le sein non-invasive et avec un coût relativement faible.

Nous présentons dans ce chapitre une nouvelle approche stochastique basée sur supports vector machine capable de classifier les tumeurs du cancer du sein. Une étude comparative en utilisant les réseaux de neurones artificiels (RNA) sera faite dans le but de valider les approches utilisées.

III.2 Modèle du sein pour la collection des données

III.2.1 Modèle de sein

Dans la littérature, plusieurs modèles du sein ont été utilisés à savoir : les modèles sphériques, cylindriques,... [65] - [66] - [67]. Le choix du modèle dépend de l'application visée. Dans notre application, nous avons utilisé un modèle donné par « S. A. AlShehri et S. Khatun », ayant la forme hémisphérique (figure III.1) et les dimensions données dans le tableau (III.1).

Partie de Modèle	Taille en (cm)
Diamètre du sein	10
Hauteur du sein	6
Epaisseur de la peau	0.2
Epaisseur de la poitrine	2

Tableau III. 1 : Dimension des paramètres du modèle

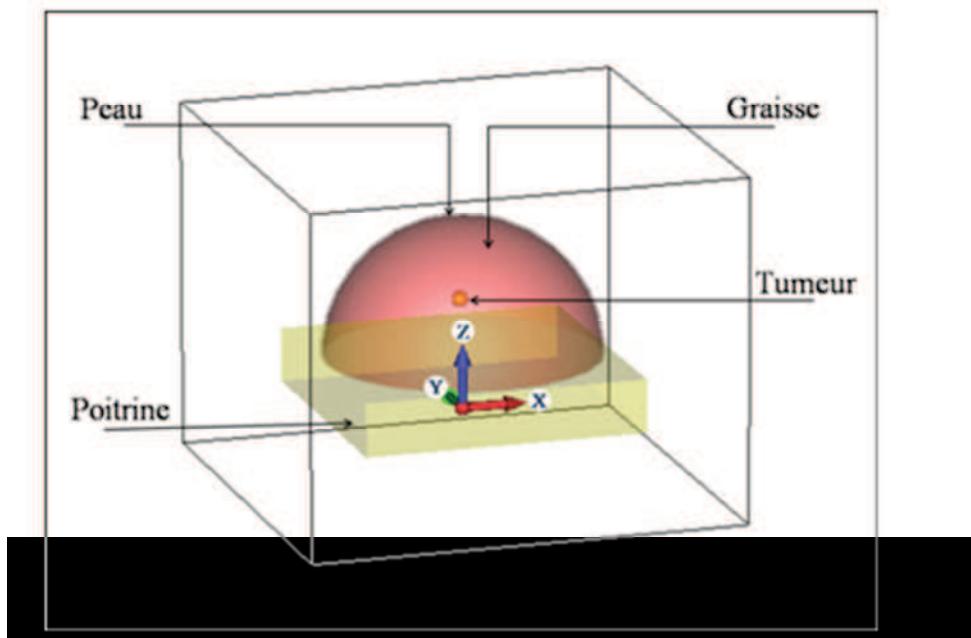


Figure III. 1 : Vue du modèle sous CST

Les propriétés diélectriques utilisées sont données dans le tableau (III.2) [67], où σ est la conductivité des tissus en (siemens / mètre) et ϵ_r est la permittivité relative.

	Conductivité σ (S/M)	Permittivité ϵ_r
Peau	1.49	37.9
Graisse	0.14	5.14
Poitrine	1.85	53.5
Tumeur	1.20	50

Tableau III. 2 : Propriétés diélectriques des paramètres du modèle

Dans la littérature [68] - [69] - [70] - [71], le rayon de la tumeur varie de 0,2 cm à 1,5 cm. Dans notre cas, nous avons pris une tumeur ayant un rayon de 0,5 cm vue qu'elle est proche de la taille minimale.

III.2.Choix de l'antenne

Dans notre application, nous avons utilisé une antenne papillon (bow-tie) représenté par la figure (III.2) pour la transmission d'une impulsion gaussienne figure (III.3) dont la fréquence centrale est de 4GHz. Cette même antenne été utilisée pour la récupération du signal diffusé (figure III.4).

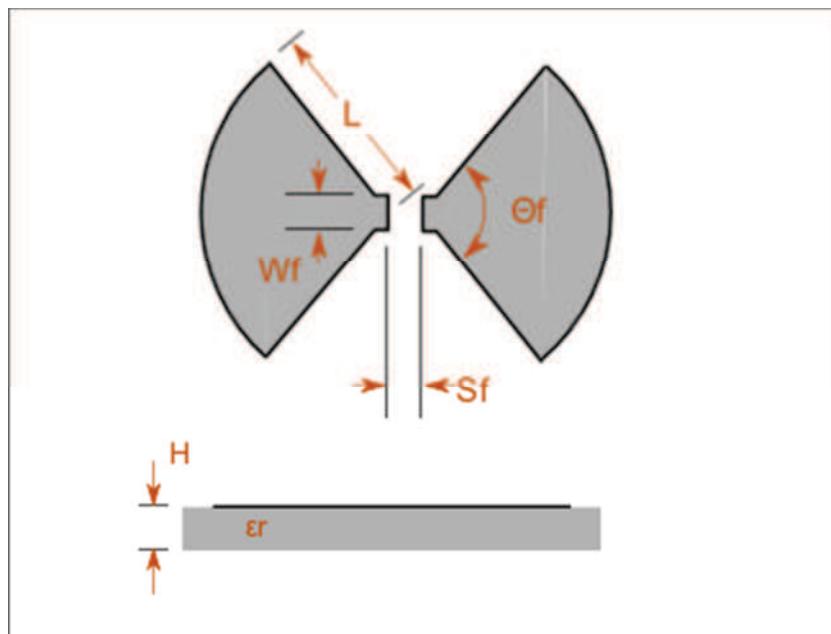


Figure III.2 : Vues de dessus et de côté d'antenne bow-tie UWB

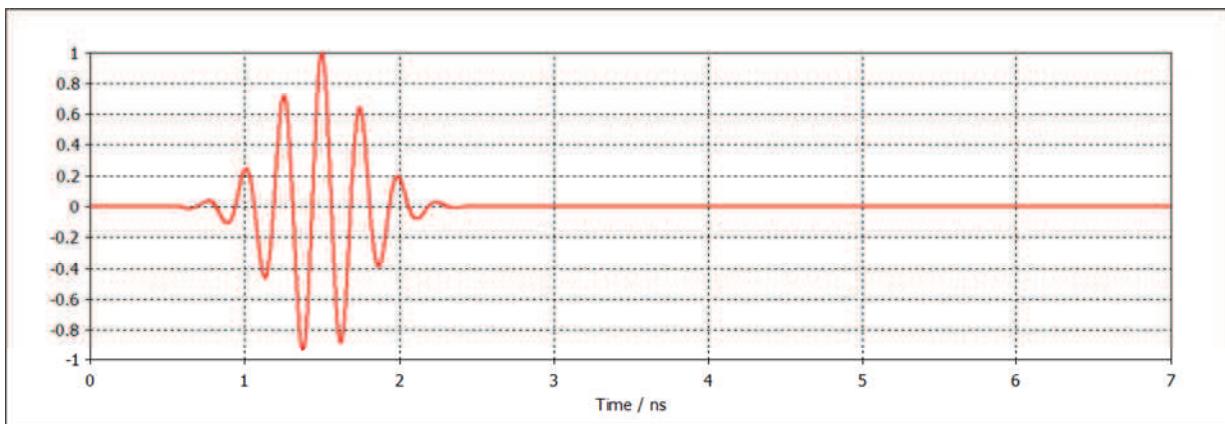


Figure III.3: Impulsion gaussienne Transmise

Ce type d'antenne est très utilisé dans notre domaine de recherche. On peut assimiler cette antenne à une antenne dipolaire dont on aurait élargi les fils dans l'optique d'augmenter la bande passante figure (III.2). Les antennes papillon présentent de nombreux avantages, tels que le faible profile, de haute efficacité de rayonnement et faciles pour la fabrication etc. Ils sont utilisés dans de nombreux domaines en raison des avantages mentionnés ci-dessus, tels que les radars à ondes pénétrantes^{[72] - [73]} et d'impulsion des antennes^{[74] - [75]}. Ces antennes ont une polarisation linéaire mais un gain et une directivité faible.

Dans ce travail, on a utilisé deux antennes papillon, dont l'une elle est utilisée pour la transmission d'impulsion gaussien et l'autre pour la récupération des données (figure III.4). Le transmetteur (Tx) et Récepteur (Rx) sont séparés par une distance de 10 cm (diamètre du sein).

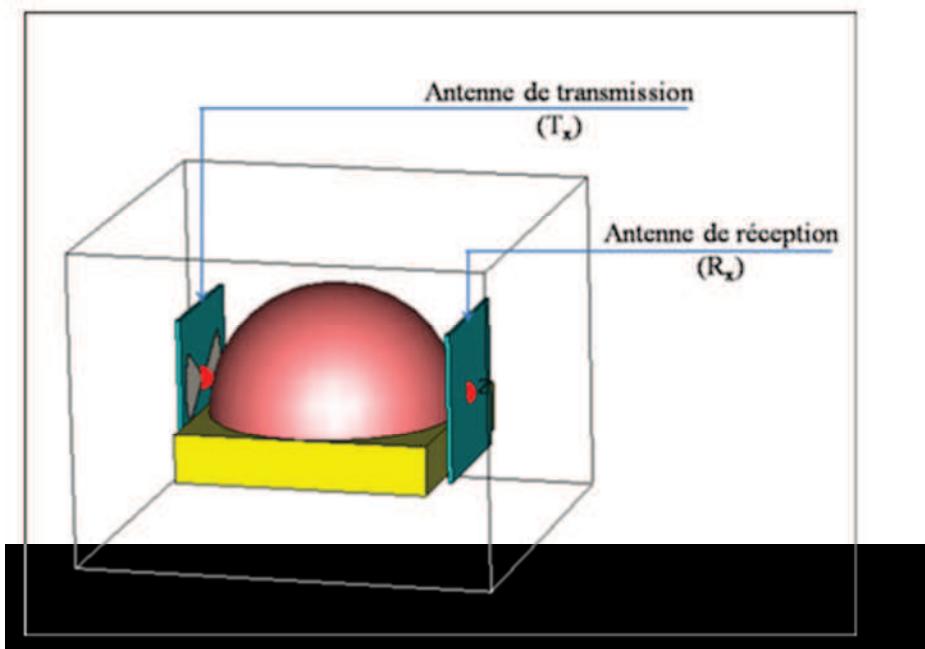


Figure III.4: Représentation des antennes dans le modèle

Les caractéristiques de coefficient de réflexion et de transmission de l'antenne sont représentées dans la figure (III.5.).

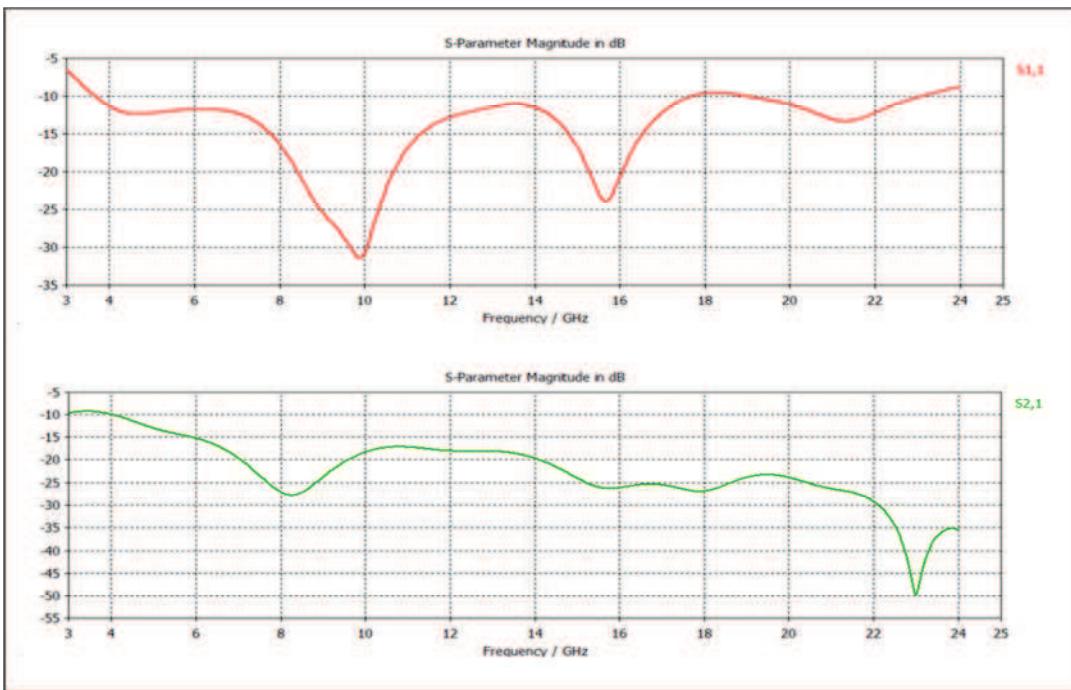


Figure III.5 : Caractéristiques S_{11} et S_{21} d'une antenne en forme bow-tie

III.3. Etapes de construction et de validation des SVM

Généralement, les étapes de construction et de validation des SVM sont réparties en trois phases :

- ❖ Choix des entrées,
- ❖ Choix des sorties,
- ❖ Tests du modèle SVM ainsi construit sur de nouveaux exemples proches des exemples d'apprentissage.

Dans nos applications, nous avons fait appel au logiciel Matlab, " Bioinformatics toolbox" afin de réaliser ces étapes [76].

La figure (III.6) représente le schéma synoptique de la phase de test du modèle construit par les supports vector machine (SVM) et SVR.

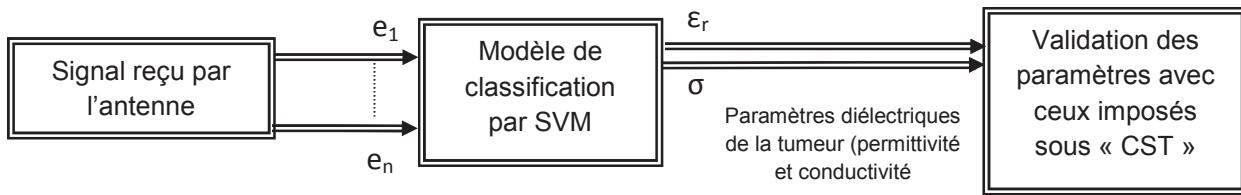


Figure III.6 : Schéma synoptique de la phase de test du modèle par SVM et SVR

III.3.1. Construction du SVM et SVR

La construction des modèles SVM et SVR se fait par un processus itératif sur les échantillons d'une base de données construite préablement. Cette base de données contient un ensemble de données (entrée/sortie) obtenues par simulation en utilisant le logiciel « CST ». Pour cela nous avons procédé de la manière suivante :

1. Placez une paire d'émetteur - récepteur à des cotés opposé du modèle du sien ;
2. Placez la tumeur dans une position fixée (dans notre cas au centre).
3. Transmettre une impulsion gaussienne d'une onde plane dans la direction de l'axe des abscisses ;
4. Recevoir le signal dans le point opposé ;
5. Changer la permittivité et la conductivité de la tumeur et répéter les étapes (3-4).

Cette procédure de génération de données a été réalisée pour 30 valeurs de permittivités et conductivités de la tumeur. En conséquence, deux groupes de signaux reçus ont été formés comme suit:

Groupe (1) : un ensemble de 20 signaux ont été utilisée pour l'apprentissage du SVM.

Groupe (2) : un ensemble de 10 signaux ont été utilisé pour la phase de test du SVM.

III.3.2. Phase de pré-traitement des données

Par ailleurs, nous avons remarqué que les signaux reçus par le récepteur contiennent un nombre d'échantillons qui peut aller de 4500 à 7200 échantillons (figure III.7). Afin de réduire ce nombre, nous avons pris uniquement un segment du signal où la signature tumorale est prépondérante. Nous avons constaté qu'une signature tumorale peut être prise dans le segment (1,25 ns et 2,13ns), cependant le nombre d'échantillons varie d'un signal à l'autre. Dans le but de fixer un pas d'échantillonnage du signal, nous avons procédé à une interpolation cubique de Hermite permettant de générer un polynôme $P(x_i)$ tout en gardant la

même allure du signal (Annexe A). Le nombre d'échantillons obtenus après interpolation avec un pas de 0.01 dans le segment (1,25 ns et 2,13ns) est de 89.

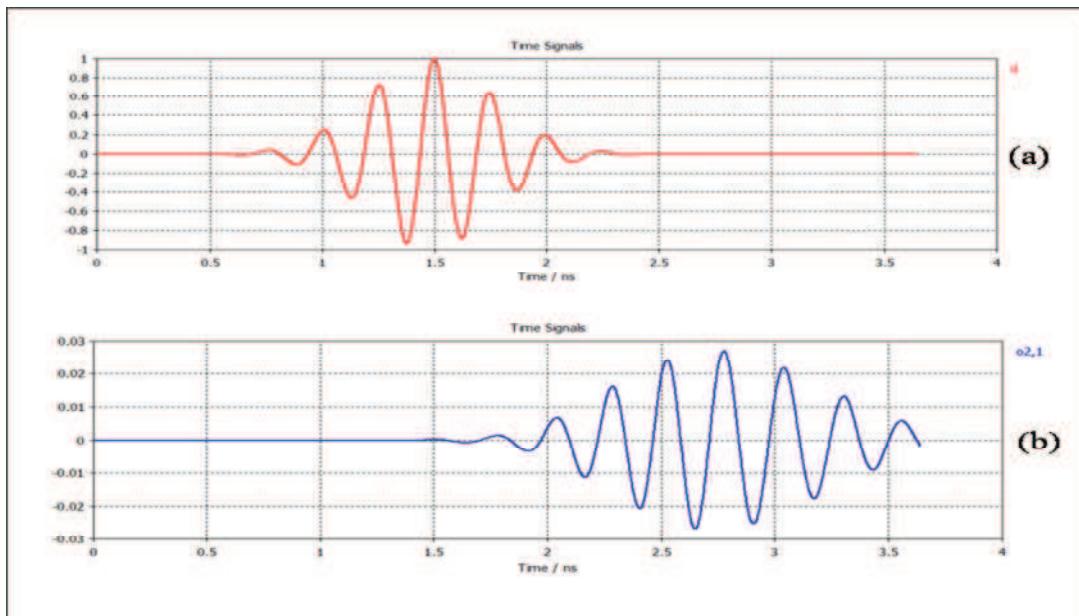


Figure III. 7 : (a) signal transmis, (b) signal reçu

III.3.3. Phase d'apprentissage

L'apprentissage vise à découvrir les relations et motifs cachées dans des collections des données brutes. Il permet également de pallier les difficultés liées à l'acquisition de connaissances auprès d'experts du domaine. Un système d'apprentissage prend en entrée un ensemble d'exemples dont il cherche une définition. Dans notre cas, on cherche à apprendre le système à classifier les tumeurs (malignes ou bénignes) à partir des signaux fournis à son entrée. Ces signaux sont présentés sous forme de paramètre (éléments du vecteur d'entrée) qui représente les signaux diffusés par la tumeur. La sortie du modèle est la permittivité relative ϵ_r dans ce cas. En fait, si $\epsilon_r < 50$ on peut qualifiée la tumeur de bénigne, sinon maligne^[77], ce qui fait correspondre deux classes à savoir (1 : maligne, 0 : bénigne).

III.3.4. Résultats de simulation et de test

Une fois la phase d'apprentissage achevée, nous avons testé les performances du SVM en utilisant le groupe 2 de test élaboré auparavant. Nous avons remarqué que le taux de classification obtenu est de 100% pour la phase de test (tableau III.3).

Groupes	Permittivité	Classe	type de la tumeur (B : bénigne, M : maligne)
Apprentissage	7.9	0	B
	9.6	0	B
	10.6	0	B
	11.6	0	B
	12.8	0	B
	15.5	0	B
	17.1	0	B
	18.9	0	B
	20.8	0	B
	22.9	0	B
	25.2	0	B
	30.5	0	B
	33.6	0	B
	45.5	0	B
	52	1	M
Test	54	1	M
	56	1	M
	58	1	M
	60	1	M
	62	1	M
	8.7	0	B
	27.7	0	B
	37.5	0	B
	7.2	0	B
	14.1	0	B

Tableau III. 3 : Résultats de simulation pour la phase d'apprentissage et de test