AIX-MARSEILLE UNIVERSITÉ



École Doctorale en Mathématiques et Informatique de Marseille

> THÈSE DE DOCTORAT en Informatique

UNIVERSITÉ TECHNIQUE DE SOFIA



École Doctorale Internationale d'Ingénierie

THÈSE DE DOCTORAT en Technique de communication et d'ordinateur

Thèse présentée pour obtenir le grade universitaire de docteur

Nicole Vesselinova CHRISTOFF

Détection et caractérisation d'attributs géométriques sur les corps rocheux du système solaire

Soutenue le 19/12/2018 devant le jury :

François COLAS	IMCCE Paris	Rapporteur
Géraldine MORIN	Université de Toulouse	Rapporteur
Ivan MOMTCHEV	Université Technique de Sofia	Examinateur
Romain RAFFIN	Aix-Marseille Université	Examinateur
Pencho VENKOV	Université Technique de Sofia	Examinateur
Agata MANOLOVA	Université Technique de Sofia	Directrice de thèse
Jean-Luc MARI	Aix-Marseille Université	Directeur de thèse
Laurent JORDA	Aix-Marseille Université	Directeur de thèse
Sylvain BOULEY	Université de Paris-Sud	Invité
Sophie VISEUR	Aix-Marseille Université	Invitée

« Une thèse, c'est un combat contre soi-même. »

Anonyme

Résumé

Détection et caractérisation d'attributs géométriques sur les corps rocheux du système solaire

L'un des défis de la science planétaire est la détermination de l'âge des terrains à la surface des différents corps célestes du système solaire. Cela sert à établir une chronologie des événements géologiques se produisant sur ces objets, donc à comprendre leurs processus de formation et d'évolution. Une approche pour la datation des surfaces du corps céleste repose sur l'analyse de la densité et de la taille des cratères d'impact. En raison de l'énorme quantité de données à traiter, des approches automatiques ont été proposées pour détecter les cratères d'impact afin de faciliter ce processus de datation. Ils utilisent généralement les valeurs de couleur des images ou les valeurs d'altitude de "modèles numériques d'élévation" (DEM). Dans cette thèse, nous proposons une nouvelle approche pour détecter les cratères, plus spécifiquement leurs bords. En effet, les cratères peuvent être caractérisés par une forme ronde qui peut être utilisée en tant que caractéristique. La méthode développée s'appuie sur une analyse de la géométrie des DEM, représentée par un maillage triangulé 3D.

L'idée principale est de combiner l'analyse de la courbure avec une classification basée sur un réseau de neurones. Cette approche comporte deux étapes principales : premièrement, chaque sommet du maillage est étiqueté avec la valeur de la courbure minimale; deuxièmement, cette carte de courbure est injectée dans un réseau de neurones pour détecter automatiquement les formes d'intérêt. Les résultats montrent que la détection des formes est plus efficace en utilisant une carte en deux dimensions s'appuyant sur le calcul d'estimateurs différentiels discrets, plutôt qu'en utilisant la valeur de l'élévation en chaque sommet. Cette approche réduit significativement le nombre de faux négatifs par rapport aux approches précédentes basées uniquement sur une information topographique. La validation de la méthode est effectuée sur des DEM de Mars, acquis par un altimètre laser à bord de la sonde spatiale "Mars Global Surveyor" de la NASA et combinés avec une base de données de cratères identifiés manuellement. Cette dernière est utilisée dans la méthode comme base d'apprentissage.

Mots clés : maillages 3D, extraction automatique de caractéristiques, courbures discrètes, réseau de neurones, détection de cratères, Mars.

Detection and characterization of geometric features on rocky surfaces on the solar system

One of the challenges of planetary science is the age determination of geological units on the surface of the different celestial bodies in the solar system. This serves to establish a chronology of the geological events occurring on these different objects, hence to understand their formation and evolution processes. An approach for dating celestial body surfaces relies on the analysis of the crater impact density and size. Due to the huge quantity of data to process, automatic approaches have been proposed for automatically detecting impact craters in order to facilitate this dating process. They generally use the color values from images or the elevation values from Digital Elevation Model (DEM). In this PhD thesis, we propose a new approach for detecting craters, more specifically their rims. Indeed, the craters can be characterized by a round shape that can be used as a feature. The developed method is based on an analysis of the DEM geometry, represented as a 3D triangulated mesh.

The main idea is to combine curvature analysis with Neural Network based classification. This approach contains two main steps : first, each vertex of the mesh is labeled with the value of the minimal curvature; second, this curvature map is injected into a neural network to automatically detect the shapes of interest. The results show that detecting forms are more efficient using a two-dimensional map based on the computation of discrete differential estimators, than by the value of the elevation at each vertex. This approach significantly reduces the number of false negatives compared to previous approaches based on topographic information only. The validation of the method is performed on DEMs of Mars, acquired by a laser altimeter aboard NASA's Mars Global Surveyor spacecraft and combined with a database of manually identified craters. The latter is used in the method as a reference-learning base.

Key words : 3D meshes, automatic feature extraction, discrete curvatures, neural networks, crater detection, Mars.

Отделяне и моделиране на геометрични характеристики на космически тела от Слънчевата система

Едно от предизвикателствата на планетарната наука е определянето на възрастта на геоложките единици на повърхността на различните небесни тела в Слънчевата система. Това служи за установяване на хронологията на геоложките събития и за разбирането на тяхното формиране и еволюционни процеси. Подходът за датиране на повърхностите на небесните тела се основава на анализ на плътността и размера на удара на кратера. Поради огромното количество данни, които трябва да бъдат обработени, са предложени автоматични подходи за автоматично откриване на ударните кратерите, за да се улесни датирането. Обикновено те използват цветовите стойности от изображенията или стойностите на коефициента на увеличение от цифров модел на издигане (Digital Elevation Model (DEM)). В настоящата дисертация предлагаме нов подход за откриване на кратери, по-специално на борда на кратера. Кратерите могат да се характеризират с кръгла форма, която може да се използва като особеност. Разработеният метод се основава на анализ на геометрията на DEM, представена като 3D триъгълна мрежа.

Основната идея е да се комбинира анализ на кривина с класификация на базата на невронни мрежи. Този подход съдържа две основни стъпки: първо, всеки връх на мрежата е означен със стойността на минималната кривина; на второ място, тази карта на кривината се подава като вход на невронната мрежа, за автоматично откриване на формите, представляващи интерес. Резултатите показват, че откриването на форми е по-ефикасно чрез използването на двумерна карта, базирана на изчисляването на дискретни диференциални оценители, стойността на височината на всеки връх. Този подход значително намалява броя на фалшиво отрицателните резултати в сравнение с други предложени подходи, основаващи се само на топографска информация. Валидацията на метода се извършва на DEM на Mapc, придобит от лазерния алтиметър на борда на космическия кораб Mars Global Surveyor на NASA и съчетан с база данни с ръчно идентифицирани кратери. Последният се използва в метода като база за референтно обучение.

Ключови думи: 3D мрежа, автоматично извличане на характеристики, дискретни криви, невронна мрежа, откриване на кратери, Марс.

Remerciements

Je tiens tout d'abord à exprimer mes plus sincères remerciements à Campus France et au gouvernement bulgare pour avoir financé cette thèse. Je voudrais également remercier Campus France Bulgarie pour son attitude professionnelle et son assistance pour chacun de mes voyages. Merci à Campus France Marseille pour son hébergement et les événements avec d'autres boursiers.

Je souhaite remercier Jean-Luc Mari et Laurent Jorda à Aix-Marseille Université, ainsi que Agata Manolova à l'Université Technique de Sofia, de m'avoir fait l'honneur d'accepter la charge d'encadrer cette thèse. J'ai apprécié leur sympathie, leur motivation, leur bonne humeur et leurs qualités pour me soutenir dans les moments difficiles dans le monde de la recherche. Je tiens à leur exprimer ma profonde gratitude pour avoir trouvé les budget pour financer toutes mes conférences et déplacements entre Bulgarie et France.

Je tiens également à remercier Sophie Viseur du laboratoire CEREGE à Aix-Marseille Université et Sylvain Bouley du laboratoire GEOPS de l'Université Paris-Sud pour leur aide et leurs conseils. Je garderai d'agréables souvenirs du temps passé ensemble en Bulgarie et en France.

Je tiens à saluer et remercier chaleureusement les membres des équipes G-Mod du LIS à Aix-Marseille Université et RKVT du Faculté de Télécommunications de l'Université Technique de Sofia pour leur accueil chaleureux. Ils m'ont aussi beaucoup aidé durant ces trois années. Je garderai un excellent souvenir de l'ambiance stimulante que nous avons entretenu en tant que collègues de travail.

À ces remerciements s'ajoutent ceux adressés à tous les doctorants (anciens et actuels). Ils ont fortement contribué à la bonne ambiance pendant cette thèse. On a partagé tout ce que comporte la vie d'un thésard et c'était magnifique!

Mes remerciements vont aussi à ma famille et mes amis qui, avec ces deux phrases « Tu es folle! », mais « Nous croyons que tu puisses le faire! », m'ont permis de ne jamais dévier de mon objectif final. Ils m'attendaient, ils étaient toujours là à chacun de mes retours.

Enfin, je tiens à souligner que rien de tout cela n'aurait été possible sans Aleksander, qui m'a soutenu et encouragé dans les moments personnels difficiles. Merci!

Table des matières

Ré	ésum	é	v
Re	emero	ciements	ix
In	Introduction générale		1
1	Les	cratères d'impact dans le système solaire	3
	1.1	Introduction	3
	1.2	Processus d'impact	3
		1.2.1 Contact et compression	5
		1.2.2 Excavation : Le cratère transitoire	5
		1.2.3 Modification et relaxation	6
	1.3	Dégradation des cratères	7
	1.4	Les différents types des cratères	8
		1.4.1 Selon la morphologie et la taille des cratères d'impact	8
			8
		Cratéres complexes	10
		1.4.2 Autre structures impacteuses	11
			11
			11
	15	Utilization des cretères d'impact	11
	1.5		12
2	Étu	de des méthodes pour la détection de cratères	13
	2.1	Données utilisées	13
	2.2	Catalogues	15
	2.3	Méthodologie pour l'étude des cratères	17
		2.3.1 Méthodes manuelles	18
		2.3.2 Méthodes automatiques	18
		Méthodes semi-automatiques	19
		Méthodes entièrement automatiques	19
	2.4	Apprentissage automatique pour la détection des cratères .	21
		Méthodes supervisées	21
		Méthodes non supervisées	23
	2.5	Bruit sur des données	24
	2.6		25
3	Extr	action de caractéristiques géométriques en 3D	27
	3.1	Les données utilisées	27
	3.2	Coordonnées géographiques et projection cartographique .	28
		3.2.1 Système de repère	28
		3.2.2 Changement de repères	29
	3.3	Courbures sur un maillage	31
		3.3.1 Description de la forme d'une surface paramétrée	31

		3.3.2 Calcul des courbures
		3.3.3 Les courbures comme descripteurs de cratères
	3.4	Prétraitement des données
		3.4.1 Réduction du bruit
		3.4.2 Valeurs RGB en niveaux de gris
		3.4.3 Génération des données d'entrée finales
	3.5	Courbes de niveau
	3.6	Algorithme basé sur un ensemble de niveaux pour l'extrac-
		tion de caractéristiques
	3.7	Comparaison avec d'autres méthodes d'extraction de carac-
		téristiques
	3.8	Conclusion
4	L'ap	prentissage automatique pour la détection de formes
	4.1	Introduction
	4.2	Choix du classifieur
		4.2.1 k plus proche voisin (kNN) \ldots
		4.2.2 Séparateurs à vastes marges (SVM)
		4.2.3 Les arbres de décision
		4.2.4 Réseau de neurones
		4.2.5 "Bagging tree"
		4.2.6 Boosting
		4.2.7 Avantages et inconvénients
	4.3	Comparaison des classifieurs selon différentes caractéristiques
		·····
	4.4	Validation croisée
	4.5	Notre choix de classifieur : le perceptron multicouche (MLP)
		4.5.1 Le perceptron multicouche
		4.5.2 Le processus d'apprentissage
	16	4.5.5 Resultats a apprentissage
	4.0	
5	Extr	raction et analyse des cratères détectés
	5.1	Détection des régions des cratères
	5.2	Extraction des bords des cratères
	5.3	Construction du catalogue Christoff <i>et al.</i>
		5.3.1 Correspondance des cratères
		5.3.2 Effacement des cratères similaires
	5.4	Analyse de catalogue Christoff <i>et al.</i>
		5.4.1 Analyse selon le diamètre
		5.4.2 Analyse selon la position

32

34

35

35

37

38

39

41

45

47

49 49

50

50

50

51

52

52

52

53

54 54

55

55

56

57

57

5	Extr	action et analyse des cratères détectés	59
	5.1	Détection des régions des cratères	59
	5.2	Extraction des bords des cratères	61
	5.3	Construction du catalogue Christoff <i>et al.</i>	62
		5.3.1 Correspondance des cratères	62
		5.3.2 Effacement des cratères similaires	63
	5.4	Analyse de catalogue Christoff <i>et al.</i>	64
		5.4.1 Analyse selon le diamètre	64
		5.4.2 Analyse selon la position	65
		5.4.3 Analyse selon le type de cratère	65
	5.5	Conclusion	68
6	Rés	ultats et validation	69
	6.1	Précision et évaluation des algorithmes de détection de cratères	69
	6.2	Évaluation	70
		6.2.1 Comparaison avec les autres catalogues	70
		6.2.2 Facteurs de qualité	72
		6.2.3 Comparaison avec d'autres méthodes	74
	6.3	Conclusion	77

7	Con	clusion et perspectives	79
	7.1	Synthèse	79
	7.2	Contributions	80
	7.3	Perspectives	80
Bi	bliog	raphie	93

Table des figures

1.1	Diagramme de formation de cratère simple et de cratère com-	4
1 2	plexe. Figure issue et modifiée de Ferrière [45]	4
1.2	et B cratère complexe où D - diamètre (bord à bord) d - pro-	
	fondeur h - hauteur des bords du cratère Illustration issue	
	de Stewart [134]	8
13	Cratère simple	9
1.5 1.4	Cratère complexe	9
1. 1 1.5	Bassin d'impact Hallas Planitia	10
1.5	Cratères secondaires	10
1.0 1.7	Cratère dégradé	11
1./		14
2.1	À gauche, une image THEMIS prise de jour. À droite, une	
	image CTX (issu de Lagain [80]).	14
2.2	Exemple de données de MOLA au format tiff, avec une réso-	
	lution de $1/128^{\circ}$ m.	14
2.3	Une partie de quatre catalogues, montrée sur une image THE-	
	MIS prise de jour avec une résolution de 100 m/pixel : (a)	
	Barlow, (b) Rodionova, (c) Robbins et (d) Salamunićcar.	16
2.4	Représentation schématique des méthodes de détection des	
	cratères	17
2.5	Les bords du cratère délimité. (a) Une région du cratère d'en-	
	trée obtenue par classification. (b) Une image binaire générée	
	par l'analyse de terrain local. (c) Une image binaire affinée.	
	(d) Un cercle trouvé dans l'image binaire affinée avec l'utili-	
	sation de CHT (issu de Di <i>et al.</i> [41]).	22
2.6	Analyse du volume de probabilité (issu de Bandeira <i>et al.</i>	
	[6]) : (a) Sélection des échantillons : (b) Filtrage selon la zone :	
	(c) Filtrage selon l'indice de circularité : (d) Détection de bord	23
27	Image binaire de la courbure topographique (à gauche) et	-0
2.7	après modifications (à droite). Issu de Bue et Stepinski [17]	24
28	Les directions longitudinale et transversale (issu de Zvl [143])	25
2.0	Les directions fongitualitale et d'ans versale (1554 de 291 [116]).	20
3.1	Repère cartésien (issu de Stout [136]).	28
3.2	Repère sphériques (issu de Stout [136]).	29
3.3	Schéma d'un cratère dans le plan cartésien	30
3.4	Courbure en un point (issu de Kudelski [76])	31
3.5	Représentation de la courbure maximale k_1 et la courbure	
	minimale k_2 qui se trouvent sur une surface (avec la permis-	
	sion de Polette <i>et al.</i> [114])	32
3.6	Les 8 catégories possibles pour un point régulier d'une sur-	
	face en considérant les signes de H et K (issu de Haala <i>et al.</i>	
	[54]).	34
	/	

3.7	Figure d'un maillage (A) avec les quatre courbures : B. Mini-	25
•	male, C. Maximale, D. Moyenne, E. Gaussienne.	35
3.8	A. Courbure minimale. B. Avec filtre d'équation 3.32 appliqué.	36
3.9	A. Histogramme de la courbure minimale. B. Histogramme	
	après le filtrage.	37
3.10	La courbure minimale en niveaux de gris	37
3.11	Les échantillons positifs sont entourés par des rectangles verts	
	et les échantillons négatifs par des rectangles rouges.	39
3.12	Fonction $q(x, y) = -x^2 - y^2$ et sa carte de niveau.	40
3.13	Représentation théorique de A) cratère simple, B) cratère com-	
0.10	nleve () cratère dégradé (issu de Christoff <i>et al</i> [28])	42
3 14	Méthode de quantification A est un cratère ancien Le cra-	14
0.14	tàra Bast angion recound par la gratàra Caratàra moyon. La	
	tere b est ancien, recoupe par le cratere C, cratere moyen. Le	10
0.15		43
3.15	Echantillonnage d'un cratere (issu de Christoff <i>et al.</i> [30]).	44
3.16	Un cratère (issu de Christoff <i>et al.</i> [30])	44
3.17	Les huit masques.	45
3.18	Les quatre courbures en niveaux de gris, avec le bruit net-	
	toyé sur le DEM redimensionné (bi-cubique) avec différentes	
	techniques appliquées pour l'extraction des caractéristiques	
	(issu et modifié de Christoff <i>et al.</i> [28])	46
4.1	Résultats avec des classifieurs différents (issu de Christoff et	
	<i>al.</i> [28])	54
4.2	L'apprentissage dans un perceptron multicouche. Le signal	
	se propage vers l'avant et l'erreur en arrière (issu de Kasabov	
	[71]).	55
	[-]).	
5.1	Carte de Mars : les zones d'apprentissage (A, B et C) et les	
	zones de validation (1, 2 et 3).	60
5.2	Résultat après CHT sur l'image $k_2 q$: a) bord ré-échantillonné.	
	b) bord délimité	62
53	Cratères «similaires» selon leur zone de chevauchement re-	0-
0.0	latif La carela rouga est "Crater 1" et la carela vert "Crater 2"	
	(issue do Christoff et al. [20])	67
F 4		02
5.4	Les crateres precis avec leurs rayons.	64
5.5	Les types des crateres (issu de Christoff <i>et al.</i> [29])	66
5.6	Les résultats de catégorisation (issu de Christoff <i>et al.</i> [29]).	67
61	Companyicon do nos anatòres avec souv du estalogue de Dob	
0.1	Comparaison de nos crateres avec ceux du catalogue de Rob-	70
	bins (a gauche) et du catalogue de Barlow (a droite)	72
6.2	Différences des diamètres des cratères catalogués / détectés,	
	respectivement pour a) site 1, b) site 2 et c) site 3. Les dia-	
	mètres des bords des cratères (km) qui sont trouvés par notre	
	approche sont tracés en bleu, les lignes jaunes sont ceux is-	
	sus du catalogue de Barlow, et en mauve ceux du catalogue	
	de Robbins.	73

6.3	Dalle 1 a), dalle 2 b) et dalle 3 c). Les cercles verts indiquent	
	tous les véritables cratères (positifs) qui correspondent au	
	moins à l'un des quatre catalogues. Les cercles bleus indiquent	
	tous les cratères étiquetés manuellement comme étant des	
	cratères. Les rectangles verts correspondent aux cratères du	
	catalogue Barlow. Les rectangles rouges sont ceux qui sont	
	omis par notre méthode. Les cercles rouges représentent les	
	mauvais cratères détectés	75
6.4	Dalle 1 a), dalle 2 b) et dalle 3 c). Les cercles verts indiquent	
	les TP. Les cercles bleus indiquent les cratères qui ne figurent	
	dans aucun des quatre catalogues. Les rectangles verts cor-	
	respondent aux cratères du catalogue de Robbins. Les rec-	
	tangles rouges sont les FN. Les cercles rouges représentent	
	les FP	76

Liste des tableaux

Comparaison des deux méthodes principales	18 20
Limites de catégorisation et valeurs d'étiquette	42
Résultats d'apprentissage du réseau de neurones	57
Informations supplémentaires sur les dalles Statistiques des cratères détectés automatiquement sur les sites 1, 2, 3, avant et après la délimitation précise du bord du	60
Cratére	65 68
Statistiques des cratères qui figurent dans le catalogue de Robbins [120]	71
Statistiques des cratères qui figurent dans le catalogue de	71
Évaluation quantitative standard de la qualité des cratères	71
d'impact avec le catalogue de Barlow	72
d'impact avec le catalogue de Robbins.	72
Comparaison avec l'approche de détection de Di [41].	77
Comparaison avec l'approche de détection de Bue et Stepinski	
[17]	77
	Comparaison des deux méthodes principales

Dédier à ...

Introduction générale

L'ÉTUDE des corps rocheux du système solaire tels que les astéroïdes, les comètes, les satellites et les planètes est fondamentale en planétologie et en géologie planétaire. Ces corps présentent de nombreux attributs géologiques communs tels que de nombreux cratères d'impact, des tranchées, des dépressions et des zones d'éboulement. Ces structures sont présentes à la surface des planètes telluriques (Mercure, Mars, Venus) et des satellites rocheux (Lune, Phobos, Deimos). De nombreuses missions spatiales, parmi lesquelles on peut citer les récentes missions DAWN de la NASA et RO-SETTA de l'agence spatiale européenne, ont révélé leur présence également à la surface de plusieurs petits corps (comètes et astéroïdes). Or, ces objets ont très peu évolué depuis leur formation dans la nébuleuse primitive qui a donné naissance aux planètes du système solaire. La compréhension des mécanismes de formation et d'évolution de ces objets passe donc par l'étude détaillée de leur surface, qui est l'un des principaux objectifs des missions spatiales lancées vers ces objets.

Les cratères d'impact sont fréquemment observés sur presque tous les corps du système solaire (planètes, satellites, astéroïdes). Ils sont dus à la collision de météorites, d'astéroïdes ou de comètes avec des objets plus massifs. L'identification de cratères d'impact à la surface des objets du système solaire ainsi que la caractérisation de leurs propriétés (profil radial et rapport profondeur sur diamètre) permettent de calculer les paramètres fondamentaux indispensables à une compréhension des mécanismes géologiques qui ont façonnés leur surface. A partir d'hypothèses sur le flux d'objets impacteurs et sur le rapport entre leur rayon et le diamètre des cratères qu'ils forment, la mesure de leur distribution en taille permet de remonter à l'âge de la surface impactée. La datation des différentes unités géologiques permet par ailleurs d'établir une chronologie relative de l'ensemble des processus géologiques qui ont façonné leur surface. La détection et la caractérisation des blocs rocheux est l'un des nouveaux champs d'étude de la surface des petits corps (astéroïdes et comètes) depuis leur détection à la surface de l'astéroïde Eros [23, 98, 78]. Ceux-ci ont une importance fondamentale puisqu'il pourrait s'agir de blocs rocheux primordiaux agrégés dans les astéroïdes et les comètes dès leur formation et révélés par les processus géologiques tels que les impacts dans le cas des astéroïdes ou le dégazage de la surface dans le cas des noyaux cométaires. La mesure de leurs propriétés, notamment de leur distribution en taille, revêt donc un intérêt fondamental.

Depuis plusieurs décennies, grâce aux missions spatiales, il est possible d'obtenir à la fois des images et des représentations géométriques et numériques de la surface des corps célestes. Ces jeux de données servent de support à la cartographie des différentes unités géologiques ainsi qu'à la datation des processus géologiques. La détection des cratères [92] et des blocs rocheux sont actuellement effectuée presque toujours manuellement. Ces méthodes deviennent inopérantes lorsque le volume de données augmente. Or, depuis les années 90, les instruments embarqués à bord des sondes spatiales interplanétaires (caméras, radars et altimètres laser) permettent une cartographie de plus en plus précise de la surface des planètes et de leurs satellites, ainsi que d'astéroïdes et de comètes. Grâce à l'utilisation de moyens informatiques de plus en plus puissants, les techniques de reconstruction 3D permettent d'extraire de ces images des modèles 3D à très haute résolution sous la forme de maillages comprenant jusqu'à plusieurs dizaines de milliards de facettes.

Plusieurs méthodes automatiques de détection de cratères ont été développées depuis une vingtaine d'années. Elles sont principalement basées sur une analyse morphologique des images et sur la reconnaissance de formes dans les images [5].

Depuis quelques années, de nouvelles méthodes utilisant directement les modèles tri-dimensionnels commencent à être développées [112].

Ce mémoire est construit autour de six chapitres. Il s'inscrit dans un cadre pluridisciplinaire entre l'astrophysique, l'apprentissage automatique et l'informatique. Dans chacun de ces domaines, des terminologies précises sont utilisées et doivent donc être définies.

Le premier chapitre définit notamment la nature des cratères d'impact, les étapes de leur formation, les differents types de cratères et les raisons qui motivent leur détection et leur caractérisation. Le chapitre II recense les différentes approches existantes dédiées à la détection des cratères sur des corps rocheux en fonction du type de données utilisé en entrée. Il est complété par un tour d'horizon des approches utilisant l'apprentissage automatique. Dans ce chapitre sont présentés les avantages et les inconvénients des techniques exposées. Le chapitre III propose une approche d'extraction des caractéristiques à partir d'un maillage et fait une comparaison avec une autre technique. Il est suivi par le chapitre IV, qui est dédié à l'apprentissage automatique pour la détection de formes. Il détaille un tour d'horizon autour du choix de classifieurs et fait une comparaison d'apprentissage par des caractéristiques différentes. Dans le chapitre V, la construction du catalogue des cratères est décrit et analysé. Le chapitre suivant montre les résultats obtenus et fait une évaluation de la méthode proposée et le chapitre VII est dédié aux conclusions et les travaux futurs.

Chapitre 1

Les cratères d'impact dans le système solaire

Sommaire

1.1	Introduction	3
1.2	Processus d'impact	
	1.2.1 Contact et compression	5
	1.2.2 Excavation : Le cratère transitoire	5
	1.2.3 Modification et relaxation	6
1.3	Dégradation des cratères	7
1.4	Les différents types des cratères	
	1.4.1 Selon la morphologie et la taille des cratères d'im-	
	pact	8
	1.4.2 Autre structures impacteuses	1
1.5	Utilisation des cratères d'impact 1	2

C^E chapitre décrit les principales étapes de formation des cratères d'impact, les différents types des cratères rencontrés sur les surface planétaires et les raisons qui motivent leur détection.

1.1 Introduction

Les cratères d'impact représentent des structures géologiques rencontrées à la surface de tous les corps rocheux du système solaire. Ils se présentent sous la forme d'une dépression de forme généralement circulaire issue d'une collision à grande vitesse entre deux objets, dont l'un des deux possède une masse suffisamment grande pour ne pas être complètement détruit par l'impact. Autour de cette dépression, une couverture d'éjecta se forme [96, 113] à partir des matériaux excavés lors de l'impact. Elle s'étend sur une surface qui dépend du diamètre de la cavité du cratère. Cette dernière est délimitée par des bords pentus. Le sol de cette cavité est composé de brèches d'impact et de roches.

1.2 Processus d'impact

Le processus d'impact dépend de l'énergie du corps impacteur, de la présence ou non d'une atmosphère autour du corps impacté, de la nature de la surface impactée et de celle de son impacteur. Si l'atmosphère de la planète est assez importante et dense, elle peut ralentir le projectile en le



FIGURE 1.1: Diagramme de formation de cratère simple et de cratère complexe. Figure issue et modifiée de Ferriere [45].

freinant et en le chauffant [15, 111]. Cela peut entraîner sa vaporisation partielle voire totale ou sa fragmentation. Dans le cas de la Terre, seuls des impacts pierreux de plus de 60 mètres de diamètre et des impacteurs de fer plus denses et plus forts de plus de 20 mètres peuvent atteindre le sol [15]. L'impact dans un environnement marin forme un cratère dans l'eau en générant de grandes vagues d'eau, voire des tsunamis. Ce mouvement d'eau peut modifier, inhiber ou même supprimer le processus de formation de cratères [151], ce qui conduit à des différences significatives entre les structures d'impacts marins et non marins [108].

La Figure 1.1 illustre le processus d'impact sur un corps rocheux tel que la surface de la planète Mars. La plupart des impacts se produisent sous un angle oblique : le pourcentage d'impacteurs qui frappent à des angles d'impact inférieurs à 60° et inférieurs à 30° sur la surface est de 75% et 25%, respectivement [130]. Lorsque l'impacteur (le projectile) atteint la surface, le processus d'impact commence. Le processus est décomposé en trois phases qui se chevauchent dans le temps : l'étape de contact et compression, l'étape d'excavation et l'étape de modification et relaxation [96]. La présentation

des trois étapes est détaillée dans les sections suivantes.

1.2.1 Contact et compression

La première phase de la formation des cratères d'impact est la phase de contact et de compression. Pendant cette étape, l'impacteur d'une vitesse supérieure à quelques km/s atteint la surface. Si la cible est une roche solide, le projectile est arrêté en une fraction de seconde. Il ne pénètre pas plus de une à deux fois son propre diamètre avant que son immense énergie cinétique ne soit transférée aux roches cibles par des ondes de choc [96, 106, 107]. L'onde de choc se propage, de façon hémisphérique, à partir du point d'impact, à la fois dans le sol cible et dans le projectile qui se trouvent tous deux compressés [96, 107] (voir Figure 1.1 A).

Le sol se fracture et les roches fondent en formant une cavité circulaire. Le matériel de la cible est mis en mouvement vers le bas et vers l'extérieur de la zone de contact (voir Figure 1.1 B). Ce déplacement se fait à la même vitesse [96]. La durée de l'étape de contact et de compression (en général quelques secondes) est déterminée par le comportement de l'onde de choc. Cette onde a été réfléchie dans le projectile à partir de l'interface projectile cible, sous la forme d'une onde de raréfaction [96] (voir Figure 1.1 B). Lorsque cette onde traverse le projectile de l'arrière vers l'avant, elle décharge le projectile des hautes pressions de choc qu'il a subi.

L'onde de choc traverse entièrement le projectile et atteint l'arrière, la surface libre de l'impacteur. La température peut atteindre plusieurs milliers de degrés dus à l'énergie apportée par l'impact. Le projectile est généralement complètement détruit ou décompressé de façon quasi-adiabatique et passe à l'état liquide et parfois à l'état gazeux [69, 96]. La fin de la phase de contact et de compression est le moment où l'onde de raréfaction atteint l'interface entre l'impacteur et la surface [47].

1.2.2 Excavation : Le cratère transitoire

L'étape de contact et compression est suivie par la phase d'excavation du matériel de la roche cible. Une partie du matériel est éjectée de la surface par l'impact et l'éjecta est propagé autour de la cavité du cratère. Le cratère transitoire représente l'agrandissement de cette cavité après le déplacement du matériel.

L'onde de choc se propage vers la profondeur et se réfléchit vers le haut sous forme de raréfaction. Dans une région de surface proche où les contraintes dans l'onde de raréfactions de tension dépassent la résistance mécanique des roches cibles, l'onde de raréfactions s'accompagne d'une fracturation et d'un éclatement de la roche cible (Figure 1.1 C). Pendant ce processus, une partie de l'énergie initiale de l'onde de choc se transforme en énergie cinétique. Le sol est fondu et se décomprime en éjectant des matériaux vers l'extérieur. Une énorme quantité de matière est expulsée vers l'extérieur à partir du point d'impact, produisant un écoulement d'excavation symétrique autour de la cavité creusée (Figure 1.1 C) lors de la phase de contact/compression (voir section 1.2.1).

Les directions du déplacement du matériel de la cible varient en fonction de l'emplacement dans les roches cibles (Figure 1.1 C). Dans les niveaux supérieurs, le matériel cible se déplace principalement vers le haut et vers l'extérieur. Aux niveaux inférieurs, la matérielle cible se déplace principalement vers le bas et vers l'extérieur. Ces mouvements produisent rapidement une dépression en forme de cuvette (cavité transitoire ou cratère transitoire) dans les roches cibles [50, 51, 94, 96] (voir Figures 1.1 D1 et D2).

Melosh [96] décrit une structure du cratère transitoire à partir d'une combinaison d'études théoriques et à partir d'études géologiques [40, 50, 52]. Le cratère transitoire idéal est une dépression en forme de cuvette avec un ligne de crète (appelé aussi bord du cratère, "rim" en anglais) structurellement soulevé (Figures 1.1 D1 et D2). Sa forme est à peu près hémisphérique, mais est en fait un paraboloïde de révolution [39].

Le diamètre d'un cratère transitoire d'au moins quelques kilomètres de diamètre est lié aux densités du projectile et de la surface [39]. La transition des cratères simples aux cratères complexes se produit à partir de 4 kilomètres environ de diamètre sur Terre, en fonction de divers paramètres relatifs à l'impacteur et à la cible [97]. Sur Mars, la transition des cratères simples à ceux ayant des caractéristiques des cratères complexes se produit sur un facteur de 3 variations de diamètre, de 3 à 10 kilomètres, une variation de l'ordre de grandeur de l'énergie cinétique de l'impact [113]. Cela signifie que dans certains cas, un cratère simple a été produit sur Mars par un impact avec 10 fois l'énergie cinétique d'un impact ailleurs qui a généré un cratère complexe.

Lorsque la résistance du sous-sol compense les forces de compression, le cratère atteint sa profondeur maximale [51, 135]. La profondeur maximale d'un cratère transitoire est d'environ un tiers de son diamètre, et cette proportion reste à peu près constante pour les cratères de taille très différente [94]. Il va bientôt subir de nouvelles modifications.

1.2.3 Modification et relaxation

La troisième étape de la formation d'un cratère d'impact est la phase de modification (voir Figures 1.1 E1 et E2). Elle commence immédiatement suivante l'étape d'excavation, qui se termine lorsque le cratère transitoire a atteint sa taille maximale. Pendant cette étape, les ondes de choc se sont décomposées en ondes élastiques à basse pression au-delà du bord du cratère. La gravité et la mécanique des roches modifient le cratère transitoire.

Durant la phase de modification se produisent les changements principaux liés aux impacts. La durée est plus longue que pour la phase d'excavation : moins d'une minute pour une petite structure jusqu'à quelques minutes pour une grande [96]. Cependant, l'étape de modification n'a pas de fin clairement déterminée. Les processus de modification de soulèvement et d'effondrement vertical de la cavité transitoire se fondent graduellement dans les processus normaux de mouvement de masse géologiques, l'enlèvement isotopique, l'érosion et la sédimentation [47].

Sur Mars, la transition des cratères simples à ceux qui ont les caractéristiques des cratères complexes se produit sur un facteur de 3 variations de diamètre, de 3 à 10 km environ, représentant plus d'une variation de l'ordre de grandeur de l'énergie cinétique de l'impact [113].

Dans le cas où la taille de la cavité transitoire ne dépasse pas un certain seuil, la résistance du matériau maintient la forme du cratère [72]. Sur la Figure 1.1 E1) est présentée l'étape de modification pour un cratère simple. L'intérieur du cratère transitoire est en fusion. Il est recouvert d'une couche de brèches (un mélange de roches brisées, fondues et cimentées) pendant le temps qu'il se refroidit. Les parois se stabilisent au rythme des effondrements et glissements de terrain qui apportent des débris à l'intérieur de la cavité et la profondeur du cratère se réduit. Ainsi le cratère final est plus large et moins profond [139].

Dans le cas où la taille de la cavité transitoire dépasse ce seuil, le cratère transitoire s'effondre et la réponse dynamique de la surface est maximale au point le plus bas de la cavité [72]. Le pic central des cratères complexes se forme par rebond élastique. Les roches comprimées par l'impact se détendent et le fond du cratère se soulève après la fin du processus d'excavation. Pendant la formation du pic central, un déficit de matière sous le cratère est créé (Figure 1.1 E2). Ensuite, les parois du cratère se tassent et forment les terrasses du cratère complexe (Figure 1.1 F2) [139].

1.3 Dégradation des cratères

Le processus de formation des cratères d'impact sur les corps rocheux du système solaire est constitué de trois phases : contact et compression (section 1.2.1), excavation (section 1.2.2) et modification et relaxation (section 1.2.3), à la fin de laquelle le cratère atteint son état final. Au cours du temps, cet état se modifie sous l'influence de l'érosion et parfois cela mène même à son effacement. Il peut également se voir complètement et brusquement détruit par des épanchements volcaniques.

Cinq processus principaux peuvent perturber la topographie des cratères [139] :

- 1. les processus fluviatiles [46]
- 2. les processus éoliens [4]
- 3. les épanchements volcaniques [58]
- 4. l'activité sismique [3]
- 5. le bombardement météoritique [1]

La topographie des cratères est associée aux dimensions principales [141]. Turtle [141] définit avec D les diamètres de cratère dans les cratères d'impact simples (A) et complexes (B). Le diamètre «final» D est également connu sous le nom de diamètre «bord à bord». Turtle [141] a défini aussi un diamètre «apparent» D_A mesuré par rapport à la surface de pré-impact. Il n'est pas montré sur la Figure 1.2, car il nécessite la reconstruction de sections érodées sur la base d'un cratère par cratère. La profondeur «finale» ddu cratère est la distance entre le haut des bords du cratère et le haut du couvercle de brèches qui recouvre le fond du cratère. La hauteur des bords du cratère est définie par h.

Sur la Figure 1.2 sont montrées les mesures principales liées aux cratères d'impact après la phase de modification (voir section 1.2.3) et après la soumission d'érosion. La Figure 1.2 A) représente un cratère simple. Le diamètre D est mesuré depuis le bord topographique d'un côté jusqu'à l'autre côté, dans des structures de cratères neufs. Le diamètre D_{er} représente le diamètre le plus susceptible d'être identifié sur la surface après l'érosion [134]. L'érosion réduit progressivement le diamètre D_{er} par rapport à D et la profondeur d_{er} par rapport à d. L'érosion réduit jusqu'à l'effacement les bords du cratère et la hauteur h tend vers 0.



FIGURE 1.2: Dégradation sous l'influence d'érosion pour A. cratère simple et B. cratère complexe où D - diamètre (bord à bord), d - profondeur, h - hauteur des bords du cratère. Illustration issue de Stewart [134].

La Figure 1.2 B) présente un cratère complexe. D_{er} est défini comme le «diamètre de l'anneau le plus à l'extérieur des failles normales concentriques (semi-)continues» (d'après Turtle *et al.* [141]). Contrairement aux cratères simples, D_{er} peut être supérieur à D, mais $d_{er} < d$ et $h \approx 0$. Cependant, la continuité des structures en anneau, les degrés d'exposition et d'érosion et la qualité de la cartographie géologique introduisent l'incertitude et la subjectivité dans la mesure de D_{er} [149].

1.4 Les différents types des cratères

Les raisons possibles des différences dans la morphologie du cratère sont étudiées par Pike [113]. L'accélération gravitationnelle et les caractéristiques de la cible affectent la morphologie des cratères d'impact à travers le système solaire.

1.4.1 Selon la morphologie et la taille des cratères d'impact

Cratères simples

Les cratères simples sont les plus fréquents à la surface des corps rocheux du système solaire. Figure 1.3 représente un tel cratère de 8 km de



FIGURE 1.3: Cratère simple.



FIGURE 1.4: Cratère complexe.



FIGURE 1.5: Bassin d'impact Hallas Planitia.

diamètre. La forme est comparable à un bol avec un profil pseudo-parabolique. Il est rare d'observer un fond plat au centre de ce type de cratère. La taille de ces cratères varie de quelques centimètres à plusieurs kilomètres de diamètre. Le rapport entre la profondeur d et le diamètre des cratères simples D est d'environ 1/5 [97]. Le fond de la cavité, en majorité, est constitué des brèches d'impact et des roches et du verre d'impact. Sous la cavité, la masse rocheuse est fragmentée [97].

Cratères complexes

Les cratères complexes sont des structures dont le centre de la cavité est occupé par un pic central autour duquel se met en place dans certains cas une unité rocheuse circulaire [96]. La Figure 1.4 représente une surface de $100 \ km^2$, au centre de laquelle est formé un cratère complexe. Il fait 35 km de diamètre. Le pic central formé lors de l'impact est au centre du cratère. Le pic central est constitué des roches déformées [96, 141]. Le fond du cratère est plat comparé au cratère simple (voir Figure 1.3). Il est composé de roches recristallisées et de brèches d'impact. Les bords du cratère complexe sont plus abrupts que ceux du cratère simple. Ces bords sont organisés en terrasses.



FIGURE 1.6: Cratères secondaires.

1.4.2 Autre structures impacteuses

Bassin d'impact

Si une météorite peut percer la croûte et provoquer des épanchements magmatiques, on parle de bassin d'impact et non plus de cratère. Le diamètre du bassin d'impact peut atteindre le millier de kilomètres. Un exemple est illustré sur la Figure 1.5. Hellas Planitia est un bassin d'impact d'environ 2100 km de diamètre et 9500 m de profondeur [138] situé au sud-est de la planète Mars (centre : $42.7^{\circ}S$ et $70, 0^{\circ}E$).

Cratères secondaires

Les cratères secondaires peuvent apparaître comme des cratères singuliers à une petite échelle, semblable à un cratère primaire avec un rayon plus petit, ou sous forme des chaînes et des groupes [55]. Sur la Figure 1.6 est représentée une chaîne de cratères secondaires. Elle représente des cratères secondaires, alignés les uns à côté des autres.

Cratères dégradés

Le cratère présenté sur la Figure 1.7 est de 18 km en diamètre. Ce cratère a subi une dégradation importante. Il ne reste aucune couverture d'éjectas ou de structure intérieure et une partie du bord du cratère a été effacée. Le bord est dégradé, le fond est plat et la profondeur du cratère est réduite.



FIGURE 1.7: Cratère dégradé.

1.5 Utilisation des cratères d'impact

Comprendre l'histoire géologique et l'évolution de l'univers commence par la connaissance de l'origine et du développement de tous les corps dans le système solaire. En l'absence d'échantillon du sol, la datation par comptage de cratères est la seule méthode pour estimer l'âge relatif des surfaces planétaires. Hartmann et Neukum *et al.* proposent de déterminer l'âge des différentes surfaces en utilisant les tailles et les fréquences des cratères d'impact [57, 59, 103, 104]. En effet, la surface des corps terrestres dans le système solaire est affectée par des astéroïdes et des météorites. Ces impacts déclenchent l'apparition de cratères de diamètres différents, en fonction de l'énergie impliquée dans la collision [47, 96].

Les cratères d'impact sont des structures d'une importance considérable. En connaissant le flux d'impact actuel et passé dans le système solaire, une étude statistique de la densité des cratères sur une surface planétaire peut nous renseigner sur l'histoire géologique de la région étudiée, la stratigraphie, l'altération et d'autres processus [56, 58, 101].

La détection des cratères d'impact est donc un problème essentiel dans la planétologie. Elle implique toujours un effort important chaque fois qu'un nouvel ensemble de données est acquis par un vaisseau spatial.

Chapitre 2

Étude des méthodes pour la détection de cratères

Sommaire

2.1	Données utilisées	13
2.2	Catalogues	15
2.3	Méthodologie pour l'étude des cratères	17
	2.3.1 Méthodes manuelles	18
	2.3.2 Méthodes automatiques	18
2.4	Apprentissage automatique pour la détection des cratères	21
2.5	Bruit sur des données	24
2.6	Conclusion	25

C E chapitre décrit un tour d'horizon des méthodes dédiées à la détection de cratères.

2.1 Données utilisées

Plus de quarante missions ont été envoyées depuis le premier survol de Mars en 1964. Ces missions spatiales ont été dédiées à l'exploration de la planète rouge. Les données acquises en orbite, grâce aux instruments présents à bord des diverses sondes, permettent d'avoir une vue globale de la composition et de la géologie de la surface martienne.

La première carte globale d'albédo de Mars est obtenu par la caméra à haute résolution MOC (Mars Orbiter Camera) à bord de la sonde Mars Global Surveyor (MGS), qui a fait une couverture globale de la surface de Mars à une résolution de 7.5km/pixel [90]. Pour la première fois, avec le spectromètre TES (Thermal Emission Spectrometer), la composition des régions sombres de Mars est captée depuis l'orbite. Les spectres de la surface de Mars sont obtenus dans l'infrarouge thermique avec une résolution spatiale d'environ 3 km/pixel [90]. L'acquisition de nouvelles bases de données d'images visibles à une résolution de plus en plus élevée est illustrée sur Mars par la cartographie complète de la planète réalisée par l'instrument Mars Express HRSC [53] et par la caméra CTX (Context Camera) à bord de Mars Reconnaissance Orbiter (MRO) [89] - 20 m / pixel et 6 m / pixel respectivement.

Un exemple de même surface, captée par la caméra THEMIS (THermal EMission Imaging System) à bord de Mars Odyssey et par la caméra visible CTX est montré sur Figure 2.1. L'instrument THEMIS peut capter des données dans le domaine visible et dans l'infrarouge. Sa résolution pour



FIGURE 2.1: À gauche, une image THEMIS prise de jour. À droite, une image CTX (issu de Lagain [80]).



FIGURE 2.2: Exemple de données de MOLA au format tiff, avec une résolution de $1/128^{\circ}$ m.

les acquisitions de jour est 100 m/pixel. Les images issues de THEMIS permettent d'observer de fins détails des cratères et d'étudier ceux de l'ordre de la centaine de mètres de diamètre [80].

Ding *et al.* [42] utilisent un ensemble de données 2D de Google Mars. Leroy *et al.* utilisent des images de Phobos, l'une des deux lunes de Mars [82]. L'image panchromatique nadir *h*0905_0000 issue de High Resolution Stereo Camera (HRSC) [102] est utilisée par Urbach et Stepinski et Liu *et al.* [142, 86]. Pour son approche Bandeira *et al.* [6] utilisent des images optiques de la surface de Mars.

Pour des objets tels que Mars, la Lune et Mercure, avec une information quasi globale de la surface, des modèles numériques d'élévation (DEM) à haute résolution sont disponibles. Concernant les surfaces planétaires, les DEM sont des descripteurs plus intéressants d'un point de vue géométrique que les images. Des données 2.5D à différentes résolutions sont utilisées. Par exemple Salamunićcar *et al.* [127] utilise des données de SELENE de résolution $1/16^{\circ}$ et Lunar Reconnaissance Orbiter ($1/512^{\circ}$). Des données DEM, générées à partir de l'altimètre laser Mars Orbiter (MOLA) avec une résolution de $1/128^{\circ}$ m à l'équateur sont utilisées par Di *et al.* [41]. Un tel exemple est montré sur la Figure 2.2.

L'état actuel de la science et de la technologie permettent de développer des approches basées sur le traitement de maillage 3D, mais malheureusement, la prévalence de telles méthodes reste rare [129].

Les relations spatiales entre les données 2D, 2.5D et 3D permettent de
lier les objets définis dans une carte sur une autre carte qui décrit les variations d'altitude. Comme résultat, tous les objets sont représentés dans un espace 3D [38].

2.2 Catalogues

Sur la Figure 2.3 sont montrés les cratères qui figurent dans quatre catalogues pour une zone particulière. La Figure 2.3 a) représente les cratères issus du catalogue de Barlow [10], à l'origine des données de Viking Orbiter. Le catalogue de Barlow est une ressource principale fiable pour l'information sur les cratères d'impact de plus de 5 km de diamètre sur Mars. Chaque entrée contient les coordonnées (latitude et longitude) du centre du cratère, la taille (diamètre et, si le cratère est elliptique, son petit diamètre et son angle d'orientation), l'unité géologique sur laquelle elle est superposée, la classe de conservation générale (les éjectas, l'absence d'éjectas mais une dégradation modérée du cratère, les «cratère fantôme»), éjecta et morphologies intérieures (le cas échéant), diamètre central du soulèvement central (le cas échéant) et éventuellement, si existe le nom du cratère.

La Figure 2.3 b) montre les cratères du catalogue de Rodionova [123]. Le catalogue contient des informations telles que les coordonnées, les diamètres, les altitudes et les caractéristiques morphologiques de tous les cratères de Mars dépassant 10 km de diamètre. Les caractéristiques morphologiques comprennent le degré de dégradation du bord, la présence de terrasses et de failles, les pics centraux, les fosses, les collines, les chaînes et les fissures, les matériaux sombres sur le fond du cratère, les éjectas fluidisés, le caractère du plancher du cratère et la surface environnante et l'albédo.

On peut noter que les cratères des catalogues de Barlow *et al.* [10] et Rodionova *et al.* [123] ne sont pas alignés avec la topographie sous-jacente. La raison est une erreur de coordonnée cartésienne (latitude / longitude) entre les images visuelles utilisées pour créer ces catalogues et des données MOLA qui sont plus précises. C'est un problème, surtout pour les petits cratères, mais les degrés de déplacement dépend de la distance de l'équateur. Plus le cratère est éloigné de l'équateur, plus le déplacement est important.

Actuellement, la base de données de cratères martiens la plus complète (contenant des cratères de plus d'un kilomètre de diamètre) a été compilée manuellement par Robbins et Hynek [120]. Cependant, les vérifications manuelles de Barlow ont montré qu'il y a potentiellement un grand nombre d'entrées erronées dans ce catalogue [13]. Pour les cratères d'un diamètre supérieur ou égal à 3 km, il existe des informations morphologiques et d'altitude. En ce qui concerne les cratères plus petits que 3 km de diamètre, seule la taille et la position sont fournis. La Figure 2.3 c) illustre les cratères du catalogue de Robbins. On n'observe pas de décalage entre les coordonnées du bord et la topographie.

Sur la Figure 2.3 d) sont montrés les cratères du catalogue de Salamuniccar *et al.* [127] pour la même zone que les Figures 2.3 a), b) et c). Le catalogue des cratères martiens est complet jusqu'à $D \ge 2$ km. Contrairement aux trois autres, qui sont collectés à la main, celui-ci est construit de manière automatique. Le catalogue est caractérisé par la fusion de quelques catalogues, dont celui de Barlow [10] et de Rodionova [123]. Il contient des



FIGURE 2.3: Une partie de quatre catalogues, montrée sur une image THEMIS prise de jour avec une résolution de 100 m/pixel : (a) Barlow, (b) Rodionova, (c) Robbins et (d) Salamunićcar.



FIGURE 2.4: Représentation schématique des méthodes de détection des cratères.

informations sur la position des centres des cratères (en latitude/ longitude), le rayon (en degrés) et un nom unique pour chaque cratère. Ce catalogue a été développé pendant des années et Salamunićcar *et al.* proposent des algorithmes différents, jusqu'aux versions finales. De même que pour le catalogue de Robbins, on ne peut pas remarquer un décalage entre les coordonnées et la topographie.

2.3 Méthodologie pour l'étude des cratères

Je vais présenter ci-dessous les méthodes de détection des cratères (voir Figure 2.4). En général, elles dépendent de la manière de traiter l'information : à la main ou automatiquement. Dans ce dernier cas, on utilise un algorithme de détection de cratères (ADC). En anglais on parle de "Crater Detection Algorithm". Je commencerai à passer en revue la littérature existant sur les ADC, à partir des données 2D et 2.5D. Je distinguerai ensuite deux catégories de méthodes : semi-automatiques et entièrement automatiques. Les premières sont basées uniquement sur le traitement d'images. Elles détectent des caractéristiques circulaires et un expert décide ensuite s'il s'agit d'un cratère ou pas. Les secondes sont basées sur la combinaison de traitement d'images et d'apprentissage automatique. C'est l'algorithme qui prend la décision si les structures sont des cratères ou non. D'un point de vue de l'apprentissage automatique, les ADCs se divisent en deux familles différentes : apprentissage supervisé et non supervisé.

	Sûre	Limitation au dia- mètre des cratères	Rapide	Facile pour faire des analyses
Manuelle	+	-	-	-
Automatique	-	+	+	+

TABLE 2.1: Comparaison des deux méthodes principales.

2.3.1 Méthodes manuelles

L'identification manuelle des cratères, appliquée aux images 2D, est une procédure bien établie et standardisée chez les géologues [22]. La méthode est très sûre (peu ou pas de faux positifs), mais prend beaucoup de temps pour les grandes surfaces qui contiennent des dizaines ou des centaines de milliers de cratères (voir Tableau 2.1).

Divers chercheurs ont collecté manuellement les coordonnées et les caractéristiques des cratères. Ces ensembles de données, certains sur toute la surface martienne [10, 79, 35, 123] et d'autres sur une région en particulier [122], reposent tous sur l'inspection visuelle des données d'imagerie.

Kuzmin [79] a cartographié des diamètres du cratère d'impact qui sont caractérisés par un ejecta fluidisé distribués sur toute la surface martienne. Le diamètre varie de 2 km à 6 km. Il estime la profondeur au sommet de la couche volatile de Mars allant de 100 m à 400 m en fonction de la distribution du diamètre initial.

L'ensemble de données le plus complet est le catalogue des grands cratères d'impact martiens [10] (ci-après appelé le catalogue Barlow). Il contient des informations sur 42283 cratères ayant un diamètre supérieur à 5 km, qui sont distribuées sur toutes la surface martienne. Construire un ensemble complet de données de cratères, comme le catalogue de Barlow, est un processus laborieux, et les résultats sont incomplets en raison des limites de la méthode de détection visuelle.

La détection manuelle par des yeux experts a été - et est toujours - la méthode couramment utilisée. La quantité de données augmente avec chaque mission spatiale. Cela impose la nécessité de ressource humaine importante pour identifier manuellement tous les cratères observés, en particulier les plus petits.

Ces lacunes augmentent la demande de techniques automatisées de détection de cratères, car elles accéléreraient considérablement le rythme de construction des ensembles de données permettraient de mettre en place un processus objectif et reproductible pour effectuer des séries d'analyses différentes. Parallèlement, le développement d'un système de navigation autonome a également provoqué un intérêt pour les algorithmes de détection automatique des cratères [82, 84].

2.3.2 Méthodes automatiques

Comme on l'a déjà mentionné, les méthodes automatiques peuvent être semi-automatiques ou entièrement automatiques.

Méthodes semi-automatiques

La méthode semi-automatique utilise un logiciel de Système d'Information Géographique (SIG) pour géoréférencer les cratères et mesurer (entre autres propriétés) leur diamètre (des outils spécifiques existent sous ArcGIS [74]).

Comme on l'a montré sur la Figure 2.3, pour certains catalogues créés à la main, il existe des distorsions des diamètres et des coordonnées après les projections. Ils peuvent être des sources d'erreurs considérables pendant les comptages des cratères pour l'estimation de l'âge des surfaces.

Pour résoudre ce problème et minimiser ces erreurs, Kneissl *et al.* ont développé une extension logicielle pour ArcMap (ArcGIS) [74]. Le logiciel a été développé pour mesurer les distributions de fréquences de tailles des cratères (en anglais : crater size-frequency distributions CSFD) sur les surfaces planétaires indépendamment des projections cartographiques des images et des données à traiter (planètes, astéroïdes, etc).

Méthodes entièrement automatiques

Les données acquises par tous les missions sont utilisées par des scientifiques pour trouver et compter les cratères d'impact sur les surfaces planétaires. Plusieurs chercheurs utilisent directement des données d'imagerie en entrée depuis quelques décennies [142]. D'autres utilisent des DEM, basés soit sur l'analyse stéréoscopique de données d'imagerie, par exemple [53], soit plus directement sur des mesures d'altimétrie laser tel que celles construites à partir de l'instrument MOLA (Mars Global Surveyor Laser Altimeter) [132, 154] ou LOLA (Lunar Orbiter Laser Altimeter) [83]. Récemment, des reconstructions 3D par photoclinométrie commencent à être utilisées [19]. Cette approche consiste à déformer un maillage jusqu'à ce que les images synthétiques de ce maillage correspondent aux images observées [19].

Ceci, associé aux nouveaux développements dans le traitement d'images, a impliqué des nouveaux efforts pour identifier un *algorithme de détection de cratères* (ADC). Le but est d'obtenir un taux de détection des cratères le plus élevé possible et un taux de fausse détection le plus bas possible. Un ADC fiable accélérerait sans aucun doute le processus de détection du cratère. Grâce à l'identification d'un grand nombre de nouveaux cratères, des analyses statistiques approfondies de la morphologie du cratère (par exemple, rapport profondeur-diamètre) seront possibles (voir Tableau 2.1). Cela donnerait à son tour accès aux éventuelles variations du taux d'impact sur la surface étudiée en plus de son âge. Dans le cas de Mars, seuls des cratères de la taille d'un kilomètre, par exemple, ont été identifiés [10].

Le développement d'un algorithme efficace de détection de cratères, exigeant un minimum d'interaction humaine, est une tâche difficile sur les plans du traitement de la géométrie et des sciences planétaires. La complexité vient du fait de la grande variété de formes de cratères, surtout en cas de chevauchement. La grande diversité des surfaces impactées joue aussi un rôle important et rend les choses plus complexes.

Les ADC automatiques ont réduit le besoin pour un opérateur de délimiter manuellement tous les cratères d'une région. De nombreuses approches [42, 93, 126, 142, 148] ont déjà été proposées pour y parvenir, mais

Auteur	Données	Surface planétaire	Apprentissage	Caractéristiques extraits	Approche utilisée (Reconnaissance du cratère)
Burl (2001) [18]	2D	Lune	Supervisé	Non	- Modèle continuellement évolutif
Duii (2001) [10]			Supervise		algorithme de correspondance (CSTM)
Leroy (2001) [82]	2D	Phobos	Non supervisé	 Détection de contours 	- Vote de tenseur
Honda (2002) [63]	2D	Lune	Non supervisé	- Détection de contours	 Transformée de Hough Algorithme génétique
Vinogradova (2002) [145]	2D	Mars	Supervisé	Non	 Modèle continuellement évolutif algorithme de correspondance
Barata (2004) [8]	2D	Mars	Non supervisé	 ACP de mesures de texture Amélioration des contours 	- Correspondance de modèle
Kim (2005) [73]	2D et 2.5D	Mars	Non supervisé	- Détection de contours - Mesure de texture - Direction de contours	 Correspondance de modèle Réseau de neurones (pour supprimer les faux positifs)
Wetzeler (2005) [150]	2D	Mars	Supervisé	Pas mentionné	- Réseau de neurones à propagation avant - Ensemble des méthodes (Bagging et AdaBoost) - Machine à vecteurs de support (SVM) - CSTM
Sawabe (2006) [128]	2D	Lune	Non supervisé	 Détection des contours "Shade and sunny" 	Multiple approches
Bandeira (2007) [6]	Images optiques	Mars	Non supervisé	 Détection des contours 	 Correspondance de modèle
Bue (2007) [17]	2.5D	Mars	Non supervisé	 Courbure topographique 	- Transformée de Hough
Martins (2009) [93]	2D	Mars	Supervisé	- Caractéristiques pseudo-Haar	- AdaBoost
Urbach (2009) [142]	Image panchromatique	Mars	Supervisé	"Highlight and shadow" algorithme	- C4.5 algorithme (arbre de décision)
Salamuniccar (2010) [125]	2.5D	Mars	Supervisé	- Détection de contours	 Machine à vecteurs de support
Ding (2011) [43]	2D	Mars	Supervisé	"Highlight and shadow" algorithme	- "Boost" algorithme - Apprentissage par transfert
Bandeira (2012) [5]	Image panchromatique	Mars	Supervisé	"Highlight and shadow" algorithme	- AdaBoost
Stepinski (2012) [133]	Image panchromatique et 2.5D	Mars	Supervisé	- Mesure de texture	- AdaBoost - C4.5 Arbre de décision
Ding (2013) [42]	2D	Mars	Supervisé	 Kanade–Lucas–Tomasi (KLT) détecteur 	 La version matricielle de SVM de type « moindres carrés »
Cohen (2014) [33]	Image panchromatique	Mars	Supervisé	- Caractéristiques pseudo-Haar	- Algorithme génétique
Di (2014) [41]	2.5D	Mars	Supervisé	 Caractéristiques pseudo-Haar Caractéristiques pseudo-Haar mise à l'échelle Motif binaire local 	- AdaBoost
Jin (2014) [68]	2D	Mars	Supervisé	 Caractéristiques pseudo-Haar 	- AdaBoost modifié
Wang (2015) [148]	Image panchromatique	Mars	Supervisé	- Mesure de texture	- Boosting (SparseBoost)
Cohen (2016) [34]	Image panchromatique	Mars	Supervisé	"Highlight and shadow" algorithme	- ConvNet
Pritchard (2016) [116]	2.5D	Lune, Mars	Supervisé	Polynôme de Tchebychev	- Arbre de décision
Palafox (2017) [110]	HiRISE	Mars	Non supervisé	Non	 Réseau neuronal convolutif (ConvNet)

TABLE 2.2: Méthodes entièrement automatiques.

elles reposent principalement sur l'utilisation d'images ou de DEM (voir la Figure 2.4 et le Tableau 2.2). Ils ne peuvent pas être distingués seulement par le type de données d'entrée, mais aussi le type de données de sortie (un "cratère simple", ou "pas un cratère", avec un ajustement de cercle, ellipse, etc.) et par le type d'approche utilisé.

Il ne suffit pas de seulement détecter les cratères. Barata *et al.* et Zhang et Jin se concentrent sur la morphologie du cratère : frais, anciens [8, 153]. D'autres auteurs ont essayé d'extraire les diamètres du cratère qui correspondent à la forme "ronde". La transformée de Hough est couramment utilisée [17, 63]. De tels ADC basés sur la détection de formes rondes ont été proposés : ajustement des cercles par Sawabe *et al.* [128] et Luo *et al.* [87]. D'autres, basés sur des données DEM, comme Salamunićcar *et al.* [127] considèrent le cratère comme ayant une forme circulaire ou ellipsoïdale.

De plus, nous pouvons distinguer les approches selon qu'elles utilisent la texture [133], la valeur de l'élévation [17] et / ou la courbure [129]. D'autres auteurs [73] combinent les données d'imagerie et les données DEM pour détecter les cratères.

En général, les ADC sont basés sur un grand nombre de méthodes pour extraction des caractéristiques, y compris la détection de contours [6, 63, 82, 125, 128] et des alternatives à la détection de contours [88], en appliquant des caractéristiques pseudo-Haar [33, 68, 93] ou bien une combinaison de techniques [41]. Stepinski *et al.* et Wang *et al.* mesurent la texture [133, 147] et Bandeira *et al.*, Cohen *et al.*, Ding *et al.* et Urbach et Stepinski [5, 34, 43, 142] utilisent la distinction des parties de surbrillance et les zones d'ombre des

cratères.

2.4 Apprentissage automatique pour la détection des cratères

Plusieurs techniques tentent d'améliorer les échantillons d'apprentissage [147] et l'apprentissage des descripteurs [43]. Récemment, dans les ADC, des approches utilisent l'apprentissage automatique. Les procédures numériques de détection entièrement automatique peuvent être distinguées en méthodes supervisées (qui nécessitent l'entrée étiquetée de données d'apprentissage) et non supervisées (entièrement autonome) (voir la Figure 2.4 et le Tableau 2.2). Elles s'appliquent aux images, DEM et moins souvent aux maillages généraux.

Le Tableau 2.2 montre une liste des méthodes entièrement automatiques, qui contiennent l'information sur le type de données utilisées, s'il s'agit d'un apprentissage supervisé ou pas et sur quelle surface l'algorithme est applicable. Pour chaque recherche, les caractéristiques extraites et les approches utilisées sont énumérées.

Méthodes supervisées

Les méthodes supervisées représentent une classification à partir des régions d'intérêts (ROI), sélectionnés sur une image [43, 68, 93] ou plus particulièrement sur des images panchromatiques [5, 34, 133, 142, 148]. De plus, Stepinski *et al.* [133] utilise des données 2.5D. Les données DEM sont aussi utilisées chez Salamunićcar et Lončarić [125], Di *et al.* [41] et Pritchard [116]. La sélection est effectuée par un opérateur pour entraîner l'apprentissage des descripteurs utilisés par le processus suivant :

- Sélection des ROIs (régions d'intérêt), constituées par des zones, qui contiennent des cratères et des zones sans cratères, souvent en utilisant un catalogue ou un spécialiste;
- 2. Spécification des descripteurs du modèle du cratère;
- Construction de la base de données avec des ROIs et apprentissage sur la base d'échantillons d'apprentissage;
- Extraction des régions candidates et classification "cratère non-cratère" en fonction de la probabilité du modèle "cratère" par l'apprentissage.

Par exemple, la plupart des méthodes supervisées appliquent une approche d'apprentissage automatique «Boosting algorithms» [43, 148]. Ils construisent de façon incrémentale un ensemble en apprentissage chaque nouvelle instance pour mettre en évidence les instances d'apprentissage qui étaient auparavant incorrectement modélisées. Un exemple typique est «AdaBoost», utilisé par Bandeira *et al.*, Di *et al.*, Martins *et al.* et Stepinski *et al.* [5, 41, 93, 133]. C'est une approche d'apprentissage automatique basée sur l'idée de créer une règle de prédiction très précise en combinant de nombreuses règles relativement faibles et inexactes. Certaines approches utilisent «AdaBoost» modifié [68].

L'arbre de décision est une méthode d'apprentissage supervisé non paramétrique utilisée pour la classification [116, 133, 142]. L'objectif est de



FIGURE 2.5: Les bords du cratère délimité. (a) Une région du cratère d'entrée obtenue par classification. (b) Une image binaire générée par l'analyse de terrain local. (c) Une image binaire affinée. (d) Un cercle trouvé dans l'image binaire affinée avec l'utilisation de CHT (issu de Di *et al.* [41]).

créer un modèle sous la forme d'une arborescence. Il le décompose et le place dans des sous-ensembles de plus en plus petits en même temps qu'un arbre de décision associé est développé progressivement. Les nœuds des feuilles représentent la classification ou la décision [117].

Ainsi, certains algorithmes de classification supervisée [42, 125] proposent une approche, qui utilise les machine à vecteurs de support, ou plus récemment apprentissage profond avec réseau neuronal convolutif (Conv Net) [34].

Les méthodes supervisées, proposées par Burl *et al.*, Vinogradova *et al.* et Wetzler *et al.* [18, 145, 150] détectent la forme ronde des cratères sur des données 2D. Deux ensembles d'images pour les deux classes et ses étiquettes sont donnés au classifieur. La détection signifie que l'algorithme entraîné détecte des cratères sur des images nouvelles et non marquées. Les auteurs de [18, 145] utilisent une technique de modèle évolutif en continu ("Continuously Scalable Template Matching Algorithme"). De plus, chez Wetzler *et al.* [150] sont testés un certain nombre d'algorithmes. Parmi eux, les machines à vecteurs de support permettent d'obtenir le meilleur résultat de détection.

La méthode de Di *et al.* [41] utilise des techniques de traitement d'images sur les données Mars DEM obtenues par l'instrument MOLA. Dans la première étape, ils utilisent un algorithme de stimulation pour détecter les entités carrées dans le DEM bidimensionnel. Trois classes de caractéristiques sont utilisées pour former l'algorithme de détection : Caractéristiques pseudo-Haar, Caractéristiques pseudo-Haar mise à l'échelle et Motif binaire local. Dans la deuxième étape, la bordure arrondie du cratère dans la région carrée est identifiée à l'aide d'une analyse locale de terrain et de la transformée circulaire de Hough (CHT) (voir Figure 2.5). Pour tester leur méthode, ils appliquent leur algorithme à trois DEM locaux de Mars, extraits des données MOLA globales. Les auteurs ont atteint un taux de détection raisonnable de 76 % et un taux de faux positifs de 13.23 %. Cependant, il n'est pas pleinement en 3D et les cratères densément répartis, se chevauchant et fortement dégradés ne sont pas détectés.



FIGURE 2.6: Analyse du volume de probabilité (issu de Bandeira *et al.* [6]) : (a) Sélection des échantillons; (b) Filtrage selon la zone; (c) Filtrage selon l'indice de circularité; (d) Détection de bord.

Méthodes non supervisées

Les méthodes non supervisées utilisent des techniques de reconnaissance de formes pour identifier les bords des cratères comme des caractéristiques circulaires ou elliptiques, telles que décrites dans [82, 128]. Pour améliorer les contours des bords, l'image originale est pré-traitée et la détection réelle est réalisée en utilisant la transformée de Hough (HT) [128]. Dans [82] comme étape de prétraitement, un algorithme de détection de bord est appliqué, les auteurs de [8] calculent des mesures de texture; et dans [73] une combinaison de mesure de texture, de détection de bord et d'analyse de direction de bord est effectuée mais pour tous les articles cités, la HT est utilisé comme méthode de détection.

En général, les ADC sont basés sur un grand nombre de méthodes, y compris la détection de cercle / ellipse [75, 140], la détection de contours [6, 63, 82, 128] voire des alternatives à la détection de contours [8, 88], la courbure topographique [17], ou la direction des contours [73] qui mesure de texture, comme dans Barata *et al.* [8].

La correspondance de modèle représente la corrélation d'un modèle et d'une image sur une fenêtre glissante. Comme résultat une image de valeurs de corrélation ou de similarité (correspondances) est produite. Cette technique est utilisée dans plusieurs algorithmes de détection [6, 8, 73]. Par exemple, la méthode de Bandeira *et al.* [6] décrit une correspondance de modèles d'images binaires. Les bords sont extraits dans une phase de sélection des candidats (voir Figure 2.6). Leurs modèles binaires sont représentés par une branche circulaire (bord) sur un fond noir, qui contient plusieurs rayons, définis par différents modèles. Bandeira *et al.* utilisent des opérateurs morphologiques pour trouver les régions maximales. L'approche semble produire de bons résultats.

La transformée de Hough est également utilisée [17, 63]. De plus, Honda *et al.* [63] utilise un algorithme génétique. Il n'est pas le seul qui s'appuie sur plusieurs approches [128]. Le vote de tenseur est appliqué par Honda *et al.* [63]. Plus récemment, l'apprentissage profond est choisi comme outil d'extraction de bord. La liste des publications liées à ADCs augmente [5].

Bue *et al.* proposent une approche non supervisée pour la détection de cratères qui utilise des données de topographie numérique au lieu d'images [17]. Ils utilisent la courbure topographique pour délimiter les bords des cratères. Les cartes des courbures sont recalculées et représentées comme une image binaire. Pour lisser les bords et éliminer les petits trous, une



FIGURE 2.7: Image binaire de la courbure topographique (à gauche) et après modifications (à droite). Issu de Bue et Stepinski [17].

fenêtre 3×3 glisse sur l'image binaire pour faire une fermeture morphologique. Puis, une squelettisation est réalisée pour la réduction de toutes les lignes à une seule épaisseur de pixel (voir Figure 2.7). La transformée de Hough est utilisée pour la détection des cratères. L'algorithme trouve de nombreux petits cratères non répertoriés dans le catalogue de Barlow, mais les cratères fortement dégradés sont difficiles à détecter. Malgré tout, le taux de détection est de 74% et la qualité des cratères détectés est de 61%.

2.5 Bruit sur des données

On a décrit les types de données et certaines de leurs sources. Pendant l'acquisition, le bruit est toujours présent. Seuls quelques algorithmes de détection automatique des cratères prennent en compte l'influence du bruit.

L'approche de Ding *et al.* [42] est basée sur le traitement d'images de Mars. Il utilise un filtre médian 3×3 pour réduire le bruit, avant l'extraction des points caractéristiques. Pedrosa *et al.* [112] proposent une approche basée sur la morphologie mathématique pendant le traitement d'images, appliquée pendant l'étape de la sélection des régions candidats. Il élimine les rayons de grain de bruit, en appliquant un filtre d'ouverture. Ce filtre dégrade les bords des grands cratères. Pour restaurer la forme originale des cratères, ils utilisent un filtre de fermeture pour relier les arêtes.

La méthode de Kim *et al.* [73], appliquée en 2D et 2.5D, semble être robuste contre le bruit et contre la variation des conditions d'éclairage d'une image à l'autre, mais il n'y a pas d'information pour les données 2.5D. Pour les données 2.5D, Bue et Stepinski [17] utilise une valeur de seuil constant pour les zones concaves. Cette constante est un compromis entre la sélectivité et la présence de bruit. L'auteur rapporte qu'il reste un peu de bruit, mais il ne perturbe pas la détection des cratères, jusqu'à ce que sa distribution spatiale ne présente pas des caractéristiques circulaires. L'auteur ne caractérise pas le bruit.

La méthode de Krøgli et Dypvik [75] utilise un "raster" de données de Fennoscandia (Terre). Dans son approche de détection automatique des contours circulaires, ils ont étudié l'effet du bruit. Ils ajoutent aléatoirement le bruit distribué (gaussien) aux modèles de données. Pour évaluer l'effet du bruit, Krøgli et Dypvik examinent le rapport, produit d'un contour placé au centre de la dépression circulaire. Le résultat de ces tests montre que si



(issu de Zyl [143]).

on augmente le bruit, le rapport diminue [75]. Sur des données MOLA, Jarmołowski et Łukasiak [67] rapporte la présence de bruit "Along-Track" et "Cross-Track".

La caméra, au bord d'instrument de télédétection, acquiert et enregistre le signal entrant de la surface. Avec l'information pour le relief de la surface, la caméra enregistre généralement une bande à la surface le long de la trajectoire de l'orbite du satellite [90]. La Figure 2.8 définit les paramètres de cette bande. La largeur représente la bande traitée ("swath en anglais") par la caméra. La direction le long de la trajectoire du satellite forme la direction longitudinale ("along-track"). La direction orthogonale à la trajectoire du satellite correspond à la direction transversale ("cross-track") [143].

2.6 Conclusion

Il existe de nombreuses approches pour la détection de cratères. Dans la reconnaissance de formes, l'apprentissage automatique est de plus en plus utilisé. Cela permet de combiner facilement expertise et automatisation. La combinaison des deux techniques (manuelle et automatique) est une bonne solution car, comme nous l'avons déjà montré, les méthodes entièrement automatiques [110] ne sont pas suffisantes pour la détection des cratères et la correction manuelle par un expert est nécessaire. Les ADC diminuent la nécessité pour un opérateur de délimiter manuellement tous les cratères dans une région, ce qui est utile pour générer des inventaires de cratères d'impact sur de vastes zones. Une inspection manuelle est toujours requise pour valider les résultats [73]. Une telle approche serait basée sur une pré-détection rapide des cratères qui servirait de base à la validation humaine.

Chapitre 3

Extraction de caractéristiques géométriques en 3D

Sommaire

3.1	Les données utilisées				
3.2	Coordonnées géographiques et projection cartographique 2				
	3.2.1	Système de repère	28		
	3.2.2	Changement de repères	29		
3.3	3 Courbures sur un maillage				
	3.3.1	Description de la forme d'une surface paramétrée	31		
	3.3.2	Calcul des courbures	32		
	3.3.3	Les courbures comme descripteurs de cratères	34		
3.4	Prétra	itement des données	35		
	3.4.1	Réduction du bruit	35		
	3.4.2	Valeurs RGB en niveaux de gris	37		
	3.4.3	Génération des données d'entrée finales	38		
3.5	Courb	es de niveau	39		
3.6	Algori	ithme basé sur un ensemble de niveaux pour l'ex-			
	tractio	n de caractéristiques	41		
3.7	Comparaison avec d'autres méthodes d'extraction de ca-				
	ractéri	stiques	45		
3.8	Conclusion 47				

C E chapitre décrit les données utilisées, leur prétraitement et l'extraction de caractéristiques géométriques associées. Ce chapitre s'appuie sur les articles : [27, 28, 30].

3.1 Les données utilisées

Dans le cadre de cette thèse, la performance de la méthode proposée sera testée sur des modèles topographiques construits à partir de l'instrument altimètre laser (MOLA) à bord de la sonde Mars Global Surveyor de la NASA [154]. L'échelle est 128 pixels par degré, ou ~ 463 m à l'équateur. Il est centré à $(0^\circ, 0^\circ)$ et représente la projection cylindrique équidistante. Les latitudes minimales et maximales sont respectivement -90° et 90° . Les longitudes minimales et maximales sont -180° et 180° également.

Les maillages triangulés 3D ne sont pas si couramment disponibles par rapport aux grilles structurées 2D issues des DEMs. Cependant, tout polyèdre peut être transformé en un maillage triangulaire sans aucune perte d'information simplement en ajoutant des arêtes sur les faces non triangulaires. Cela signifie que l'approche proposée peut être appliquée à tous les types de données topographiques. En outre, les maillages utilisés dans ce travail de recherche ont une distribution homogène identique des sommets sur leur surface.

Les données sources sont stockées dans des fichiers de type OBJ ou PLY, c'est-à-dire des index sur les faces polygonales. Les fichiers contiennent une liste de sommets de coordonnées tridimensionnelles (x, y, z) et les index sur les faces. Pour faciliter le traitement, la surface de Mars est divisée en "dalles" situées entre deux longitudes et latitudes limites.

3.2 Coordonnées géographiques et projection cartographique

3.2.1 Système de repère

Il est nécessaire à définir les trois repères que l'on va utiliser et de pouvoir passer de l'un à l'autre. La surface martienne est caractérisée par sa position dans l'espace et ses représentations bidimensionnelles : en fonction de coordonnées cartésiennes planaires et en fonction de représentation comme pour une image 2D :

- Coordonnées cartésiennes tridimensionnelles (3D)
- Coordonnées cylindriques (2D cylindriques)
- Coordonnées bidimensionnelles (pixels 2D)

Le modèle géométrique nous permet de re-projeter le point (O, x, y, z) dans le plan bidimensionnel pour réaliser notre localisation spatiale.

Coordonnées cartésiennes tridimensionnelles (3D) Les coordonnées cartésiennes tridimensionnelles R^3 sont associées à la planète Mars que nous observons. Ces coordonnées sont également utilisées pour caractériser la forme tridimensionnelle de l'objet. La position de chaque point M est définie par les coordonnées cartésiennes x, y et z.



FIGURE 3.1: Repère cartésien (issu de Stout [136]).

Le repère cartésien est définie par les axes Ox, Oy, Oz. On peut décrire un point M ou un vecteur $\overrightarrow{A}(M)$ par les vecteurs unitaires $(\widehat{e_x}, \widehat{e_y}, \widehat{e_z})$. En cas de point, ils sont respectivement dans les directions des axes du système, avec un point d'origine O (voir Figure 3.1 a)). Pour le repérage du vecteur, le point d'origine est situé au point M (voir Figure 3.1 b)).

Un point *M* de l'espace est repéré par les trois composantes du vecteur \overrightarrow{r} , qui relie l'origine *O* à *M* :

$$\overrightarrow{r}(x,y,z) = x\widehat{e_x} + y\widehat{e_y} + z\widehat{e_z}$$
(3.1)

dont les composants x et y montrent les coordonnées de la projection du point M dans le plan xOy, définie par le point M'.

Coordonnées sphérique Les coordonnées dans ce repère sont définies en radians.



FIGURE 3.2: Repère sphériques (issu de Stout [136]).

Un point M de l'espace est repéré par les trois composants :

- 1. z sa côte par rapport au plan xOy
- 2. ρ , qui représente le rayon du cylindre
- 3. ϕ , l'angle entre l'axe Ox et la droite OM'

et est projeté dans le plan xOy au point M' (voir Figure 3.2).

Coordonnées bidimensionnelles dans une image 2D Les coordonnées dans ce repère sont définies en pixel.

3.2.2 Changement de repères

Au cours de ce travail, il est nécessaire de passer d'un repère à un autre. On présente les trois principaux passages des systèmes de coordonnées :

De la 3D vers la 2D sphérique On considère un repère orthonormé pour travailler sur des distances et des angles :

$$OM = \sqrt{x^2 + y^2 + z^2}$$
(3.2)

$$\phi = \arcsin\left(\frac{z}{OM}\right) \tag{3.3}$$

$$\lambda = \operatorname{arctg}\left(\frac{y}{x}\right) \tag{3.4}$$



FIGURE 3.3: Schéma d'un cratère dans le plan cartésien

De la 2D sphérique vers la 3D On peut définir les distances entre le point projeté M' et les points O et M par :

$$OM' = R.cos\phi \tag{3.5}$$

$$MM' = R.sin\phi \tag{3.6}$$

En utilisant le rayon R et les angles λ et ϕ , les coordonnées 3D sont :

$$x = R.\cos\phi.\cos\lambda \tag{3.7}$$

$$y = R.cos\phi.sin\lambda \tag{3.8}$$

$$z = R.sin\phi \tag{3.9}$$

où les dimensions des angles peuvent être déterminées, en utilisant les relations entre les distances :

$$\sin\phi = \frac{z}{OM} \tag{3.10}$$

$$\cos\lambda = \frac{x}{OM} \tag{3.11}$$

Pour chaque cratère, on calcule les coordonnées cartésiennes du centre, obtenues par la carte bidimensionnelle 2D, comme montré sur la Figure 3.3. Le diamètre *D* en [*km*] en longitude λ et latitude ϕ est calculé à partir des équations suivantes :

$$D'[^{\circ}] = \frac{D[km]}{\delta P[km]} \times 128^{\circ}/px$$
(3.12)

$$t_i = \delta t \times i \qquad \qquad t = 0 - 360^\circ \qquad (3.13)$$

Les coordonnées du centre (X, Y) de chaque cratère peuvent être définies par :

$$X_{e,i} = D_x \cos\frac{t_i}{2} \tag{3.14}$$

$$Y_{e,i} = D_y \sin \frac{t_i}{2} \tag{3.15}$$



FIGURE 3.4: Courbure en un point (issu de Kudelski [76]).

3.3 Courbures sur un maillage

En géométrie différentielle, on vise à étudier les propriétés locales d'une géométrie par le biais d'outils issus du calcul différentiel. Nous nous intéressons plus particulièrement à la notion de courbure. D'un point de vue géométrique, la courbure en un point d'une courbe contenue dans un plan est définie comme étant l'inverse du rayon du cercle approchant au mieux la courbe au voisinage de ce point. Cet estimateur peut être étendu à un point d'une surface plongée dans l'espace tridimensionnel.

Cet indicateur est largement utilisé dans le traitement des maillages car il offre des informations sur la concavité ou la convexité d'une surface le long de ses directions principales.

Les courbures peuvent être calculées sur un DEM (donc en 2,5D), mais, afin de proposer un algorithme générique, nous avons choisi de travailler sur des maillages 3D à valeurs multiples. Ainsi, l'approche peut être appliquée non seulement sur de gros corps comme Mars (où les DEM sont disponibles) mais aussi sur des corps plus petits (comme Vesta, Ceres ou Lutetia) où des acquisitions 3D complètes sont disponibles, et pas seulement les élévations codées dans des DEM.

3.3.1 Description de la forme d'une surface paramétrée

Les formules et les notions utilisées dans cette thèse pour le calcul des courbures sont issus de Carmo [21], Mynatt *et al.* [100] et Kudelski [76]. On donne des brèves explications sur les applications aux surfaces. Dans ce manuscrit, comme dans Mynatt *et al.* [100], on utilise la description donnée par Lipschultz [85] pour décrire la forme d'une surface paramétrée dans l'espace euclidien 3D. Deux équations aux dérivées partielles sont utilisées, elles représentent les premières et deuxièmes formes fondamentales. Elles caractérisent la façon de déterminer les surfaces, les longueurs et les angles d'une surface.

Dans le plan \mathbb{R}^2 , la courbure en un point p d'une courbe contenue est définie comme étant l'inverse du rayon r du cercle (le cercle osculateur) approchant au mieux la courbe (la courbe plane) au voisinage de ce point



FIGURE 3.5: Représentation de la courbure maximale k_1 et la courbure minimale k_2 qui se trouvent sur une surface (avec la permission de Polette *et al.* [114]).

(voir Figure 3.4). La courbure k en un point est définie par :

$$k = \frac{1}{r} \tag{3.16}$$

Dans \mathbb{R}^3 , une définition plus complète pour une surface est nécessaire. On écrit une surface *S* par la forme paramétrique :

$$S(u,v) = \begin{pmatrix} x(u,v) \\ y(u,v) \\ z(u,v) \end{pmatrix}, (u,v) \in \mathbb{R}^2$$
(3.17)

Le vecteur normal N est défini par la relation :

$$N = \frac{X_u \times X_v}{\|X_u \times X_v\|},\tag{3.18}$$

où \times dénote le produit vectoriel et X_u et X_v sont les dérivées partielles de S :

$$X_u = \frac{\partial S\left(u, v\right)}{\partial u} \tag{3.19}$$

$$X_{v} = \frac{\partial S\left(u,v\right)}{\partial v} \tag{3.20}$$

3.3.2 Calcul des courbures

Considérons le plan, défini par (p, N, v) (voir Figure 3.5), une surface triangulée avec un point d'intersection p de deux cercles osculateurs, avec une normale N et vecteur tangent v. La courbure maximale (k_1) et la courbure minimale (k_2) , où $k_1 \ge k_2$, sont représentées en rouge et en bleu. Les

directions principales sont orthogonales et associé aux valeurs de k_1 et k_2 , avec :

$$k_1 = \frac{1}{r_1}$$
(3.21)

$$k_2 = \frac{1}{r_2}$$
(3.22)

Il existe une autre méthode pour calculer les courbures et les directions principales. La courbure normale *k*, pour toutes les directions en tout point de la surface, est définie par le rapport :

$$k = \frac{\mathbb{II}}{\mathbb{I}} \tag{3.23}$$

où I et II sont respectivement les premières et secondes formes fondamentales. Le vecteur différentiel $d\vec{x}$ est la tangente à la surface avec une direction définie par du et dv. Mynatt *et al.* [100] décrivent la première forme fondamentale (I), d'après Lipschultz [85], comme le produit intérieur de $d\vec{x}$ avec lui-même :

$$I = d\vec{x} \cdot d\vec{x} = (\vec{x_u} du + \vec{x_v} dv) \cdot (\vec{x_u} du + \vec{x_v} dv)$$

= $(\vec{x_u} \cdot \vec{x_u}) du^2 + 2(\vec{x_u} \cdot \vec{x_v}) du dv + (\vec{x_v} \cdot \vec{x_v}) dv^2$
= $E du^2 + 2F du dv + G dv^2$ (3.24)

Pour le calcul des angles, des distances et des surfaces sur la surface, Mynatt *et al.* [100] utilise les premiers coefficients fondamentaux (E, F et G). A partir des dérivées partielles, la matrice de la première forme fondamentale est obtenue :

$$\mathbb{I} = \begin{pmatrix} E & F \\ F & G \end{pmatrix}$$
(3.25)

La seconde forme fondamentale quantifie la vitesse spatiale de changement de \hat{N} dans chaque direction, déterminée par $d\vec{x}$ en un point de la surface, en prenant le produit interne des deux vecteurs différentiels :

$$\begin{split} \mathbb{II} &= -d\vec{x} \cdot d\hat{N} = -(\vec{x_u}du + \vec{x_v}dv) \cdot (\hat{N}_u du + \hat{N}_v dv) \\ &= -(\vec{x_u} \cdot \hat{N}_u)du^2 - 2(\vec{x_u} \cdot \hat{N}_v + \vec{x_v} \cdot \hat{N}_u)dudv - (\vec{x_v} \cdot \hat{N}_v)dv^2 \\ &= Ldu^2 + 2Mdudv + Ndv^2 \end{split}$$
(3.26)

Pour la caractérisation locale de la surface (i.e. elliptique, parabolique ou hyperbolique) Mynatt *et al.* [100] utilise les deuxièmes coefficients fondamentaux (L, M et N). De même, la matrice \mathbb{II} est caractérisée par :

$$\mathbb{II} = \begin{pmatrix} L & M \\ M & N \end{pmatrix} \tag{3.27}$$

A partir des premières et secondes formes fondamentales, on peut calculer l'endomorphisme de Weingarten L :

$$L = \mathbb{I}^{-1}\mathbb{I}\mathbb{I} \tag{3.28}$$

	K < 0 : hyperbolic	K = 0 : parabolic/planar	K > 0 : elliptic
H < 0			
H = 0			not possible
H > 0			

FIGURE 3.6: Les 8 catégories possibles pour un point régulier d'une surface en considérant les signes de H et K (issu de Haala *et al.* [54]).

où \mathbb{I}^{-1} est l'inverse de la matrice \mathbb{I} .

Il est à noter, que les valeurs propres de L définissent les grandeurs des courbures principales au point donné et les vecteurs propres de L caractérisent les orientations des courbures principales dans le plan.

3.3.3 Les courbures comme descripteurs de cratères

On peut décrire la forme et l'orientation en un point de la surface par la courbure gaussienne K et la courbure moyenne H. Ces deux estimateurs sont liés à la courbure minimale k_2 et à la courbure maximale k_1 par les relations suivantes :

$$H = \frac{(k_1 + k_2)}{2} \tag{3.29}$$

$$K = k_1 \cdot k_2 \tag{3.30}$$

En un point régulier d'une surface, il est possible de définir les huit types de formes, en utilisant les signes de H et K [76]. La représentation de ces formes est illustrée par la Figure 3.6.

Un cratère d'impact est caractérisé par une dépression circulaire entourée d'une crête circulaire. Ceci est dû à la dissipation isotrope de l'énergie cinétique de l'impacteur sur la surface pendant l'impact. En utilisant les courbures moyenne et gaussienne, la zone de dépression peut être délimitée par des sommets avec H > 0 et K > 0 tandis que la zone de crête est représentée par des sommets avec H < 0 et $K \approx 0$ (voir la Figure 3.6). Cependant, ce critère ne prend pas en compte le chevauchement possible de multiples cratères.

La Figure 3.7 montre une zone d'un échantillon de Mars. Les zones convexes sont en bleu (H < 0), les zones relativement plates sont en vert ($H \approx 0, K \approx 0$) et les zones concaves sont en jaune (H > 0). Le signe de H caractérise la convexité locale ou la concavité de la surface. Ceci souligne



FIGURE 3.7: Figure d'un maillage (A) avec les quatre courbures : B. Minimale, C. Maximale, D. Moyenne, E. Gaussienne.

l'importance de la perception pour un algorithme de détection basé sur la géométrie.

En ce qui concerne la forme géométrique des cratères d'impact, caractérisée par une dépression circulaire entourée d'une crête circulaire, la courbure minimale est choisie comme un bon descripteur de la concavité locale de la surface. L'argumentation de ce choix est visible sur la Figure 3.7. La courbure Moyenne et la courbure Gaussienne ne permettent visiblement pas d'obtenir un contour continu. Des tests décrits dans la section 4.3 démontrent l'utilité de la courbure Minimale en comparaison de la courbure Maximale.

3.4 Prétraitement des données

3.4.1 Réduction du bruit

Sur la Figure 3.8 A. est montré la courbure minimale, calculée en chacun des points sur une dalle de Mars. Les "Along-Tracks" et "Cross-Tracks" sont bien visibles. Ce bruit est étudié par Jarmołowski et Łukasiak [67].

Gauthier *et al.* [49] proposent une méthode efficace pour la caractérisation du bruit dans les processus 3D. Elle est basée sur l'analyse de l'histogramme de courbure des sphères "bruyantes". Cette approche est basée sur l'hypothèse que Mars est une sphère parfaite, donc nous pouvons déduire que des arguments similaires peuvent être appliqués.

Pendant l'une de ses expériences, Gauthier *et al.* [49] ajoutent de bruit gaussien à une géométrie de sphère idéale. L'exemple conduit à des histogrammes gaussiens des valeurs de courbure. L'histogramme des valeurs de courbure minimales calculées à partir de l'ensemble de données (voir Figure 3.8 A), est illustré sur la Figure 3.9 A. Après inspection visuelle, on



FIGURE 3.8: A. Courbure minimale. B. Avec filtre d'équation 3.32 appliqué.

peut conclure que la forme de l'histogramme est semblable à celle d'un histogramme gaussien. Cela peut signifier que pendant l'acquisition de la topographie de Mars, le bruit pourrait être considéré comme gaussien. Sur la base de cette hypothèse, on le nettoie. La première phase se détermine par le calcul du quantile de la distribution.

On cherche les valeurs qui sont situées en dehors de l'intervalle dans seulement 5% des cas (voir Rees [118]) :

$$X_{1,2} = \mu \pm 1.96 * \sigma \tag{3.31}$$

où X_1 et X_2 sont les variables aléatoires standards, μ est la valeur moyenne et σ est l'écart-type. Pour préserver les relations spatiales de voisinage des valeurs de la courbure, on applique un filtre 2D linéaire adaptatif :

$$\frac{1}{8} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$
(3.32)



FIGURE 3.9: A. Histogramme de la courbure minimale. B. Histogramme après le filtrage.

sur les valeurs qui entrent dans les intervalles :

$$] - \infty, X_1] \cup [X_2, +\infty[$$
 (3.33)

Le résultat de ce filtrage est présenté sur la Figure 3.8 B. L'histogramme de la distribution de ces valeurs est montré sur la Figure 3.9 B. On peut observer que les "Along-Tracks" et "Cross-Tracks" sont supprimés sans endommager la topographie de la surface.

3.4.2 Valeurs RGB en niveaux de gris

De nombreuses méthodes utilisent une image de 8 octets, comme équivalente à une image en niveaux de gris, pour l'extraction de caractéristiques de couleur. C'est également une pratique bien établie au lieu d'utiliser les espaces de couleur d'origine [70]. Il faut remarquer que les niveaux de gris représentent des couleurs RGB converties en faible gamme de valeurs.



FIGURE 3.10: La courbure minimale en niveaux de gris.

Pour chaque dalle, on trouve les valeurs minimale $k_{2,min}$ et maximale $k_{2,max}$. À partir de ces repères, on a converti les valeurs des courbures minimales k_2 en une information de niveaux de gris quantifiée entre 0 et 255.

La matrice de sortie est constituée des valeurs de la courbure en niveaux de gris " $k_{2,G}$ ", ayant les mêmes dimensions (X et Y) que le maillage. L'image de sortie de cette étape de prétraitement, appelées ci-après "image", est utilisée comme entrée pour notre algorithme.

Sur la Figure 3.10 est présenté un exemple de cette quantification. La courbure minimale d'entrée est la même que celle présentée sur la Figure 3.8 B. Les valeurs 0 et 255 sont déterminés par $k_{2,min}$ et $k_{2,max}$ respectivement. La carte d'information pour $k_{2,G}$ est de dimension 3200×1665 pixels.

3.4.3 Génération des données d'entrée finales

L'étape suivante se caractérise par la préparation d'un ensemble de données pour l'apprentissage d'un classifieur. Dans cette étude, on se sert de la version 2 du catalogue de Barlow (Barlow Northern Hemisphere Database) [9, 12] pour l'extraction des échantillons contenant des cratères. Les coordonnées des pixels du centre (X et Y) des cratères sur $k_{2,G}$ sont calculées à partir de leurs coordonnées géographiques (latitude et longitude) et stockées dans un fichier de données. Leur diamètre est converti de kilomètres à pixels en utilisant l'échelle de l'image $k_{2,G}$.

Pour chacun des $k_{2,G}$ qui vont servir comme zones d'apprentissage, on prend les régions autour tous les centres des cratères du catalogue de Barlow. On a déjà expliqué dans le Chapitre 2 Section 2.2 le décalage des coordonnées X et Y avec la topographie Martienne, qui diffère d'un point de vue de la direction et de la taille.

Les bords des cratères sont retracés par des cercles de rayon égal au rayon de chaque cratère de l'ensemble de données [9]. Ils sont positionnés aux coordonnées qui figurent dans le catalogue. Pour corriger toutes les erreurs possibles dans la position du cratère dans le catalogue Barlow, on a repositionné manuellement tous les cercles déjà calculés.

Comme l'échelle minimale des pixels de l'équateur est de 463 m, une sous-fenêtre d'environ 1.5 fois le diamètre du cratère (voir Di *et al.* [41]) plus 5 pixels sont automatiquement extraits autour de la position centrale. La constante de 5 pixels est choisie pour gérer les erreurs de la position des centres (elle avait été choisie en fonction de la différence moyenne calculée suite aux corrections manuelles). Cela permet d'obtenir une région d'intérêt, représentant une sous-fenêtre, contenant le cratère, y compris les éventuels éjectas qui l'entourent. On utilise aussi ces cratères, dont les centres sont dans les limites de la région traitée, mais ils ne sont pas entièrement inclus dans la région.

De plus, de cette région d'intérêt, qui contient un cratère (cercle au milieu), on a automatiquement sélectionné un nombre équivalent de régions sans cratère. Les tailles sont comprises entre 2 et 100 pixels. Ces régions d'intérêt sans cratère ne chevauchent pas les sous-fenêtres contenant un cratère. On a vérifié manuellement que les régions "sans cratères" ne contiennent pas de cratères. Pour rappel, le catalogue de Barlow contient des cratères de plus de 5 km de diamètre. On a supprimé ces régions et en a sélectionné de nouveau.

Sur la Figure 3.11 sont illustrées les positions et tailles de tous les échantillons extraits d'une des dalles d'apprentissage. Les rectangles verts sont les sous-fenêtres autour de tous les cratères dans le catalogue de Barlow dans la zone donnée. Les rectangles rouges sont les sous-fenêtres qui servent



FIGURE 3.11: Les échantillons positifs sont entourés par des rectangles verts et les échantillons négatifs par des rectangles rouges.

d'échantillons négatifs. Elle contient 455 échantillons positifs et le même nombre d'échantillons négatifs.

Toutes les sous-fenêtres sont rééchantillonnées pour créer une fenêtre normalisée de 20×20 pixels (comme Di *et al.* [41]).

L'interpolation bilinéaire est appliquée par Bandeira *et al.* [5] et Di *et al.* [41]. D'autres auteurs, comme De Croon et Izzo ou Szabo *et al.* par exemple, utilisent l'interpolation bicubique [36, 137]. En général, elle est considérée comme une interpolation préférable, parce qu'elle conserve un semblant de netteté. Elle bénéficie de plus de données pour le calcul.

Nous avons essayé les deux techniques. Dans nos publications [26, 27, 28], l'interpolation bilinéaire été appliquée. Les résultats dans nos publications [29, 30] sont obtenus avec une l'interpolation bicubique. On pense que cette dernière est préférable, car les courbes sont plus lisses si on applique l'interpolation bilinéaire. Elle introduit moins d'artefacts qui se distinguent comme dégradant la qualité.

En comparaison avec l'interpolation bilinéaire, l'interpolation bicubique augmente un peu la complexité du calcul. Toutes les sous-fenêtres sont rééchantillonnées par interpolation bicubique pour créer une fenêtre normalisée de 20×20 pixels, ci-après appelée *Area*.

3.5 Courbes de niveau

Notre approche est inspirée par les courbes de niveau, qui sont utilisées pour afficher les points du relief à la même altitude sur une carte. Il s'agit des lignes de niveau, qui sont utilisées comme descripteur général des courbes.

D'après la définition donnée par Boyd [16], on peut tracer des lignes de niveaux sur un terrain bidimensionnel avec des courbes unidimensionnelles sur lesquelles la quantité tracée q est une constante. Ces courbes sont définies par :

$$q(x,y) = q_j, \qquad j = 1, 2, ..., N_c$$
 (3.34)



FIGURE 3.12: Fonction $q(x, y) = -x^2 - y^2$ et sa carte de niveau.

où N_c est le nombre de contours tracés. Ces courbes de constante q sont appelées les "contours" de q, les "courbes de niveau" de q ou les "surfaces de niveau" de q.

Considérons la fonction :

$$q(x,y) = -x^2 - y^2 \tag{3.35}$$

Elle a la forme d'un paraboloïde.

La Figure 3.12 a) illustre le graphique de la fonction de l'équation 3.35. Le graphique est coupé avec j plans de niveau écartés régulièrement, parallèles au plan xy. On peut considérer ces plans comme les lieux où z est égal à la sortie donnée [-4, 4]. La Figure 3.12 b) marque les endroits où le graphique est coupé par les plans. Ces lignes (Figure 3.12 c)) sont projetées sur le plan xy et la carte de niveau à laquelle ils correspondent (Figure 3.12 d)). On peut noter que les cercles ne sont pas régulièrement échelonnés. En s'éloignant de l'origine, la hauteur du graphique augmente plus rapidement. En conséquent, l'augmentation de la hauteur nécessite un plus petit pas par rapport à l'origine dans l'espace d'entrée [2].

En utilisant les courbes de niveau, il est possible de connaître l'altitude, le profil d'un itinéraire, les cols, les sommets et les crêtes. Plus les courbes de niveau sont rapprochées, plus la pente est raide. De même, plus les courbes sont espacées, plus la pente est douce.

L'équidistance est la distance verticale entre deux courbes de niveau. Si on connaît l'équidistance, il est possible de calculer la hauteur. Le point le plus haut est situé au dessus de la dernière courbe et avant la création d'une nouvelle courbe. Le point le plus bas est situé au dessus de la première courbe. L'absence de courbe signifie une zone plate. Un itinéraire sur une même courbe signifie un déplacement à flanc.

Pour connaître le sens des courbes, on peut utiliser les points d'altitude, et l'on peut repérer sur la carte deux points d'altitude : le plus haut et le point le plus bas.

3.6 Algorithme basé sur un ensemble de niveaux pour l'extraction de caractéristiques

Les courbes de niveau tracent des lignes sur une carte à travers tous les points ayant la même valeur d'une quantité mesurable. Autrement dit, ils utilisent l'altitude comme descripteur. On propose de sélectionner des niveaux sur une surface et de construire une carte des caractéristiques, à travers tous les points ayant les mêmes attributs. En d'autres termes, on va utiliser la géométrie de la surface pour décrire les points (caractéristiques) ayant la même description en utilisant une relation mesurable. Il est à remarquer que la méthode d'interpolation influe la fiabilité de représentation des lignes de niveau, la représentation des creux, des pics et de la pente, respectivement [37].

Dans le Chapitre 1, on a décrit les étapes de formation, à travers lesquelles le cratère atteint sa morphologie terminale. Dans le cas général, le cratère est caractérisé par un bord, l'éjecta autour de la cavité, le fond, et il peut contenir un pic central au milieu. La structure intérieure peut être lisse (pour les cratères simples) ou avec des terrasses (pour les cratères complexes). Le diamètre, la profondeur et la hauteur des bords dépendent de l'énergie d'impact et la structure de la surface.

L'approche ici proposée est basée sur l'idée que le bord du cratère est plus haut que le fond du cratère. Le rapport entre la hauteur du bord et la profondeur du cratère est plus grand pour un cratère frais que pour un cratère dégradé. La méthode devrait répondre aux exigences suivantes :

- Être applicable pour des tailles différentes (rayons différents);
- Être indépendante de la rotation et de l'angle d'observation;
- Être convenable pour des morphologies différentes;
- Être adaptée selon la topographie régionale (y compris le cas de chevauchement);
- Être capable de prédire l'état de dégradation.

On s'appuie sur l'hypothèse de Ponti *et al.* [115] disant que l'utilisation d'un nombre de couleurs réduites avec une méthode de quantification appropriée réduira considérablement la dimensionnalité, tout en améliorant ou en maintenant la précision d'un système de classification.

On propose la séparation de chaque Area (définie dans la section 3.4.3) en six zones, à la manière d'une réduction de dimensionnalité, tout en conservant les informations descriptives du cratère. Comme repères, on propose d'utiliser les valeurs : minimale (Min), maximale (Max) et moyenne (Mean), calculées pour les Areas respectives.



FIGURE 3.13: Représentation théorique de A) cratère simple, B) cratère complexe, C) cratère dégradé (issu de Christoff *et al.* [28]).

Chacune de ces zones (dans le cas général) va contenir l'information suivante :

- Intervalle 1 : la crête du bord du cratère
- Intervalle 2 : la hauteur du cratère et l'apparition des terrasses
- Intervalle 3 : la hauteur de l'éjecta
- Intervalle 4 : la surface avant l'impact
- Intervalle 5 : la surface vers le fond
- Intervalle 6 : le fond du cratère

Sur la Figure 3.13 on présente schématiquement un cratère simple (A.), un cratère complexe (B.) et un cratère en état de dégradation (C.), considéré être valable pour les deux cas A et B. Les positions des valeurs minimale, maximale et moyenne sont représentées par des points. Les intersections montrent les positions des contraintes calculées. Les domaines (zones) sont les plages des valeurs en elle-même.

Considérons les échantillons positifs, sélectionnés dans la section 3.4.3. Ils contiennent au moins un cratère. La méthode énoncée est une méthode de quantification rapide qui consiste à marquer les valeurs de chaque *Area*, comme décrit dans le Tableau 3.1.

Domaines	étiquette
[Min, Mean(1 - (1 - P))/2]	1
$\left((Mean(1-P) + Min)/2, Mean(1-P) \right)$	2
(Mean(1-P), Mean]	3
(Mean, Mean(1+P)]	4
(Mean(1+P), Max(1-P)]	5
(Max(1-P), Max]	6

TABLE 3.1: Limites de catégorisation et valeurs d'étiquette.

On utilise un paramètre P, comme une valeur de seuil entre les niveaux, pour définir toutes les valeurs dans le domaine $k_{2,G}$. On a essayé avec plusieurs valeurs différentes. Nous proposons que P soit entre 2% et 3% et inférieur à 4%. À ce titre, les informations significatives du descripteur de cratère (telles que le fond du cratère, le bord, etc.) seront décrites. Tous nos résultats sont obtenus avec P = 3%.

La Figure 3.14 est une représentation de la méthode pour l'extraction de caractéristiques basée sur six niveaux, en utilisant les règles de quantification de la Table 3.1. Le cratère A est un très gros cratère ancien, avec



FIGURE 3.14: Méthode de quantification. A est un cratère ancien. Le cratère B est ancien, recoupé par le cratère C, cratère moyen. Le cratère C est frais, chevauché par un cratère moyen.

un rayon de 55 px. Le cratère B est ancien, recoupé par le cratère C, cratère moyen avec un rayon de 16 px. Le cratère C est frais, chevauché, de cratère moyen de 8 px de rayon.

Les Figure 3.14 A), B) et C) décrivent le voisinage des cratères présents dans le catalogue de Barlow. Le cratère dans 3.14 B) peut être vu sur le côté inférieur droit de 3.14 A) et celui dans 3.14 C) sur le côté supérieur droit de 3.14 B). Les Figures 3.14 a), c) et e) représentent la fenêtre redimensionnée 20×20 autour de chaque cratère et les b), d) et f) sont leurs représentations respectives en niveaux de gris. Les six images sont quantifiées selon l'approche proposée en utilisant différentes valeurs pour le paramètre *P*. L'objectif est de montrer que quelle que soit la taille de la fenêtre, l'algorithme proposé traite bien la visualisation des caractéristiques du cratère. Il est à noter que P = 2% convient aux cratères d'un diamètre supérieur à 6 km. Après des tests expérimentaux, nous avons obtenu les meilleurs résultats avec des valeurs différentes, pour détecter de plus petits cratères avec des diamètres inférieurs à 6 km, avec P = 3%.

L'information extraite d'*Area*, qui peut contenir (ou non) des cratères, ci-après appelée *newArea*, est un ensemble d'attributs de catégorisation, obtenu avec la procédure de quantification expliquée dans le Tableau 3.1. Les valeurs révèlent certaines caractéristiques d'une zone particulière de l'image, telles que les bords ou les changements de texture.

La Figure 3.15 (a) est la représentation en maillage d'un voisinage d'un petit cratère frais, après le filtre de bruit; la Figure 3.15 (b) est la courbure minimale calculée k_2 du même cratère. En appliquant la répartition des niveaux sur la Figure 3.15 (c2) (ou la Figure 3.15 (c1)), la transition entre le bord du cratère et la surface de Mars, la surface elle-même et la pente du fond sont perdues. La Figure 3.15 (d2) est la représentation en niveaux de



FIGURE 3.15: Échantillonnage d'un cratère (issu de Christoff *et al.* [30]).



FIGURE 3.16: Un cratère (issu de Christoff et al. [30]).

gris des valeurs sur la Figure 3.15 (c2). Ayant ces hypothèses, nous ne pouvons pas donner de valeurs absolues.

Pour les cratères frais (voir la Figure 3.16), le bord reçoit les valeurs 1 (bleu) et 2 (vert). Du haut vers l'intérieur du cratère : les valeurs de 3 (jaune) et 4 (orange) représentent la transition entre le bord et le fond du cratère. Les valeurs de 5 (rouge) sont la transition de la surface avant l'impact et du fond du cratère, ayant la valeur 6 (violet). Du haut vers l'extérieur : le 3 représente aussi la transition entre le bord et la surface. L'éjecta des cratères est étiqueté avec 4 (orange). La surface est étiquetée avec 5. En ce qui concerne le rapport rayon/profondeur du cratère et le niveau de dégradation, il est possible que certaines valeurs ne soient pas présentes.



FIGURE 3.17: Les huit masques.

3.7 Comparaison avec d'autres méthodes d'extraction de caractéristiques

Le but principal de l'application des algorithmes d'extraction des contours est la possibilité d'extraire des segments qui suivent réellement la topographie du terrain. Cela définit l'extraction de la forme exacte des caractéristiques géomorphologiques.

On a déjà énuméré les méthodes les plus utilisées : elles appliquent des techniques d'extraction des contours comme pseudo-Haar, Motif binaire local, etc.

Motif binaire local (LBP) L'opérateur de motif binaire local est un descripteur de texture popularisé par Ojala *et al.* [105], qui extrait des informations sur le contraste local. La texture est définie en fonction des variations spatiales de l'intensité en pixels d'une image. Il est utilisé pour encoder des informations sur les motifs réguliers présents dans une image, ainsi que pour détecter les bords et les arêtes.

Le LBP est considéré dans une fenêtre flottante de 3×3 pixels sur chaque échantillon précédemment extrait. Chaque image est donc représentée sous une forme vectorielle.

Caractéristiques pseudo-Haar La seconde information extraite de chaque échantillon de courbure en niveaux de gris redimensionné (ci-après C_g) est un ensemble de caractéristiques pseudo-Haar. Cette méthode, définie par Viola et Jones [146], est un algorithme rapide utilisé pour détecter les contours d'intérêt dans une image. Il consiste à appliquer des masques composés de rectangles (voir Figure 3.17), séparant l'image en zones claires et sombres. Tout d'abord, les valeurs de pixels à l'intérieur de la zone noire sont ajoutées ensemble; les valeurs dans la zone blanche sont alors additionnées. La différence d'intensité de pixel (courbures) dans ces deux zones est calculée à chaque étape. La méthode génère des vecteurs pour chaque image. Nous utilisons huit masques différents de 2, 3 et 4 rectangles. Ces masques ont la taille de 20×20 pixels. Les valeurs révèlent certaines caractéristiques d'une zone particulière de l'image, telles que des arêtes ou des changements de texture.

Comme cela se produit dans les cascades LBP, les classifieurs faibles deviennent des classifieurs puissants lorsqu'ils sont organisés en séquence dans une cascade de type pseudo-Haar. Ils sont capables de détecter des structures malgré l'éclairage, la couleur ou les variations d'échelle.

Caractéristiques pseudo-Haar mise à l'échelle Les caractéristiques sont introduites par Di *et al.* [41]. La fonction de pseudo-Haar mise à l'échelle implique un redimensionnement des fonctionnalités de type pseudo-Haar



FIGURE 3.18: Les quatre courbures en niveaux de gris, avec le bruit nettoyé sur le DEM redimensionné (bi-cubique) avec différentes techniques appliquées pour l'extraction des caractéristiques (issu et modifié de Christoff *et al.* [28]).

précédemment calculées. Formellement, si nous calculons un vecteur de caractéristiques de pseudo-Haar Hr_{ps} sur un échantillon de taille w redimensionné à une nouvelle taille s (ici 20 pixels) pour une résolution de courbure principale à l'équateur c, alors les caractéristiques de Haar sur C_g sont calculées par :

$$C_q = Hr_{ps.} (cw/s)^{-1} \tag{3.36}$$

Cette formule découle du principe suivant : les plus gros cratères sont généralement plus profonds que les plus petits. La caractéristique de pseudo-Haar mise à l'échelle est calculée par division de la valeur de caractéristique d'Haar avec un coefficient de résolution de l'échantillon pour ajuster les caractéristiques du type Haar des grands cratères et des petits cratères à la même échelle. Le nombre de fonctionnalités d'Haar ressemblant à celles de tous les échantillons redimensionnés est identique à celui des fonctionnalités de type Haar.

Sur la Figure 3.18 est montré un même cratère, représenté par ces quatre courbures : en colonnes k_1 , k_2 , H, K. La cinquième colonne représente l'entrée du DEM. Toutes les entrées sont d'abord visualisées en dimension

réelle (ligne 1). Les courbures sont recalculées en niveaux de gris, avec le bruit préalablement réduit. Toutes les entrées sont ensuite redimensionnées en fenêtre 20×20 (ligne 2). La ligne 3 montre les caractéristiques pseudo-Haar aux premiers niveaux de décomposition. Les caractéristiques pseudo-Haar mises à l'échelle, proposées par Di *et al.* [41] sont montrées sur la ligne 4. Dans la cinquième ligne sont calculés les motifs binaires locaux. Dans la ligne 6, les résultats sont présentés après extraction des caractéristiques, comme proposé dans ce manuscrit.

Notre descripteur donne comme résultat un petit vecteur (400 éléments), au lieu de 13397 (4714 + 4714 + 3969) éléments tels que proposé dans les travaux de Di *et al.* [41], et qui consistent à combiner les caractéristiques de Haar, de pseudo-Haar mis à l'échelle et LBP en un vecteur (les auteurs testent l'algorithme sur un DEM). Grâce à cela, l'algorithme proposé utilise 5 fois moins de temps de traitement et beaucoup moins d'espace disque.

3.8 Conclusion

Dans ce chapitre, nous avons proposé une méthode qui utilise la géométrie comme outil d'analyse des surfaces et de description des formes. Cette approche aide à suivre des formes qui changent de topologie et aussi de caractériser des formes qui varient dans le temps, comme les cratères d'impact. Il s'agit d'une approche qui extrait des caractéristiques assez descriptives pour décrire les différents groupes de cratères sur une topographie variée.

Comme avantage du modèle, on peut ajouter que les calculs de courbures sur des maillages 3D sont effectués sur une grille cartésienne sans avoir besoin de passer d'un repère à un autre. Comme avantage, le temps du traitement des courbures n'affecte pas le temps d'exécution [109].

Chapitre 4

L'apprentissage automatique pour la détection de formes

Sommaire

4.1	Introduction				
4.2	Choix du classifieur				
	4.2.1 k plus proche voisin (kNN) $\ldots \ldots \ldots 50$)			
	4.2.2 Séparateurs à vastes marges (SVM) 50)			
	4.2.3 Les arbres de décision	l			
	4.2.4 Réseau de neurones	2			
	4.2.5 "Bagging tree"	2			
	4.2.6 Boosting	2			
	4.2.7 Avantages et inconvénients	3			
4.3	Comparaison des classifieurs selon différentes caracté-				
	ristiques	1			
4.4	Validation croisée 54				
4.5	Notre choix de classifieur : le perceptron multicouche				
	(MLP)	5			
	4.5.1 Le perceptron multicouche	5			
	4.5.2 Le processus d'apprentissage	5			
	4.5.3 Résultats d'apprentissage	7			
4.6	Conclusion	7			

C E chapitre présente différentes approches d'apprentissage automatique pour la détection des cratères. Pour la construction de ce chapitre, on s'appuie sur nos articles : [26, 28].

4.1 Introduction

L'apprentissage automatique est utilisé pour l'amélioration d'extraction des échantillons [147] ou bien pour l'apprentissage des descripteurs [43]. Les méthodes de détection des cratères se divisent en méthodes supervisées et non supervisées. En général, en classification, on a un ensemble de classes prédéfinies ("cratère" ou "pas cratère") et on cherche à savoir à quelle classe une nouvelle région appartient. Le partitionnement de données tente de regrouper un ensemble de régions et de déterminer s'il existe une relation entre elles.

De nombreux algorithmes sont construits en utilisant des méthodes supervisées, tels que Wang *et al.* [147], Stepinski *et al.* [133], Ding *et al.* [42] et Vinogradova *et al.* [145]. Les algorithmes "boosting" sont couramment utilisés [147], et "AdaBoost" en particulier [133]. Les techniques comme les séparateurs à vastes marges et les modèles évolutifs sont préférés par Ding *et al.* [42] et Vinogradova *et al.* [145] respectivement. Cependant, Kim *et al.* [73] propose une approche basée sur l'utilisation de réseau de neurones.

Les algorithmes de détection des cratères comme ceux proposés par Bandeira *et al.* [6], Barata *et al.* [8], Bue et Stepinski [17] et Leroy*et al.* [82] utilisent des approches non supervisées. Ce type de détection de cratères est préféré pendant la navigation repérée optiquement [25]. Bue et al. utilisent la transformée de Hough [17] et Barata *et al.*[8] mesurent la texture. D'autres techniques sont aussi pratiquées, telles que la correspondance de modèle [6] ou le "tensor voting" [82].

4.2 Choix du classifieur

Les différents modèles de classification sont utilisés pour attribuer des échantillons de la surface en jeu à un groupe discret ou de classes basées sur un ensemble spécifique de caractéristiques.

4.2.1 *k* plus proche voisin (kNN)

La règle du plus proche voisin (1NN) et son extension de K plus proches voisins (KNN) est l'une des approches les plus simples pour effectuer une classification [44]. Chaque nouvel échantillon est attribué à la classe de son plus proche voisin. Soient $x = x_1, ..., x_N$ de n échantillons et $y = y_1, ..., y_c$ l'ensemble des classes possibles. Soit X' un échantillon qui appartient à xet voisin le plus proche d'un point de test X. La règle du plus proche voisin dit que le point X sera affecté à la même classe que celle du point X'. La règle des K plus proches voisins peut être, estimée ainsi :

$$P(y_i|X) = \frac{K_j}{K} \tag{4.1}$$

Elle peut être interprétée comme une règle de recherche du maximum de la probabilité a posteriori, où K_i représente le nombre de voisins, qui appartiennent à la classe *i* parmi les *K* plus proches voisins. La valeur de *K* est liée aux densités : une valeur élevée sera adaptée aux densités faibles et une valeur faible, aux densités élevées [91].

4.2.2 Séparateurs à vastes marges (SVM)

Les machines à vecteurs de support peuvent être appliquées dans deux cas (de données) : séparable et non séparable. Le cas non séparable se diffère dans le but de justifier les erreurs de classement si c'est nécessaire. Puis, on introduit un coût supplémentaire. Les ξ_i , i = 1, ..., l sont introduites pour se faire des variables d'écart positives. La distance des points à la marge
sont chiffré par ces variables [91] :

$$\beta^T x_i + \beta_0 \ge 1 - \xi_i, \quad \text{pour} \quad y_i = 1 \tag{4.2}$$

$$\beta^T x_i + \beta_0 \le -1 - \xi_i, \quad \text{pour} \quad y_i = -1$$
(4.3)

$$\xi_i \ge 0 \quad \forall i \tag{4.4}$$

Au cas où un échantillon serait mal classé, la variable ξi correspondante sera supérieure à un et il y aura une erreur.

Les machines à vecteurs de support sont des méthodes de noyaux. Ces fonctions sont utilisées avec les SVM pour permettre au classifieur de traiter des classes non linéairement séparables. En modifiant l'espace de caractéristiques, ces fonctions essayent de le transformer en un problème linéairement séparable. Plusieurs types de noyaux sont couramment utilisés [31] : linéaire, gaussien ("Radial Basis Function (RBF)") [64], quadratique, cosinus et etc.

4.2.3 Les arbres de décision

Les arbres de décision sont une méthode d'apprentissage supervisée non paramétrique utilisée pour la classification et la régression. L'objectif est de créer un modèle sous la forme d'une arborescence. Elle décompose un ensemble de données en sous-ensembles de plus en plus petits en même temps qu'un arbre de décision associé est développé de manière incrémentale. Le résultat final est un arbre avec des nœuds de décision et des nœuds feuilles. Un nœud de décision a deux branches ou plus. Chaque échantillon appartient à cette classe la plus courante dans la région. Le taux d'erreur de classification est le nombre d'observations dans cette région qui n'appartiennent pas à la classe. Cependant, il n'est pas suffisamment sensible, et une mesure statistique de la dispersion de la distribution des échantillons dans la classe (tel que le coefficient de Gini par exemple) et l'entropie croisée sont préférables [66].

Le coefficient de Gini mesure de la variance totale dans les classes K :

$$G = \sum_{k=1}^{K} \hat{p}_{mk} \left(1 - \hat{p}_{mk} \right)$$
(4.5)

où \hat{p}_{mk} est la proportion d'observations d'apprentissage dans la région m, appartenant à la classe K. Le nœud contient principalement des échantillons d'une seule classe si G a une petite valeur. L'entropie croisée est très similaire au coefficient de Gini [66], donnée par l'équation suivante :

$$D = \sum_{k=1}^{K} \hat{p}_{mk} log \hat{p}_{mk}$$
(4.6)

Une petite valeur de D indique que le nœud contient principalement des échantillons d'une seule classe.

4.2.4 Réseau de neurones

L'expression mathématique d'un neurone biologique est un neurone formel, qui a une ou plusieurs entrées. Pendant la transmission d'informations dans un réseau, un neurone formel reçoit et transmet cette information à d'autres neurones du réseau. À l'entrée il reçoit une somme pondérée des signaux qui proviennent de différents neurones connectés à lui. À la sortie, le neurone formel envoie un signal, qui représente une fonction de cette somme pondérée [71] :

$$y_i = f\left(\sum_{i=1}^N w_{ij} x_i\right) \tag{4.7}$$

Le neurone formel *j* reçois les signaux x_i , (i = 1...N) des neurones *i*. Les poids entre ces neurones d'entrée et de sortie (i et j) sont marqués par w_{ij} . Pour chaque neurone *j*, la sortie y_j est calculée. La fonction *f* est une fonction d'activation. Les réseaux avec plus d'une couche se caractérisent par : 1) les neurones entrées, qui sont liés aux attributs (données); 2) les neurones de sortie, qui sont liées à une classe donnée et 3) les neurones cachés, qui sont entre 1) et 2).

4.2.5 "Bagging tree"

Le "bagging tree" est une méthode d'ensemble, qui construit plusieurs arbres de décision en rééchantillonnant à répétition l'ensemble des données avec remplacement. C'est une approche qui réduit la variance des arbres de décision. Un tel modèle à faible variance peut être obtenu en déterminant les nombre d'arbres $\hat{f}^1(x), \hat{f}^2(x), ..., \hat{f}^C(x)$ d'un ensemble *C*, et leur moyenne [66] :

$$\hat{f}_{avg}(x) = \frac{1}{C} \sum_{c=1}^{C} \hat{f}^{c}(x)$$
(4.8)

Si on a un seul ensemble de données il est possible de dupliquer les échantillons de données. Cela génère différents ensembles de données C. On utilise le c^{me} ensemble d'échantillons bootstrap pour effectuer la formation et obtenir $\hat{f}^{*c}(x)$:

$$\hat{f}_{bag}(x) = \frac{1}{C} \sum_{c=1}^{C} \hat{f}^{*c}(x)$$
(4.9)

4.2.6 Boosting

Le boosting est analogue de "bagging tree", sauf que les arbres sont obtenus de manière séquentielle : chaque arbre est construit en utilisant des informations d'arbres déjà assemblés. L'ajustement d'arbre est effectué sur une version modifiée d'ensemble de données d'origine. Le boosting combine un grand nombre d'arbres $\hat{f}^1(x)$, $\hat{f}^2(x)$, ..., $\hat{f}^L(x)$ d'ensembles L.

L'ajustement d'un arbre se fait par les résidus actuels r_i et pas avec le résultat Y comme réponse : $\hat{f}(x) = 0$ avec $r_i = y_i$ pour tous les *i* dans

l'ensemble d'apprentissage. Chaque arbre l = 1, 2, ..., L peut être constitué que quelques nœuds terminaux, déterminés par le paramètre d de l'algorithme [66]. L'ajustement du \hat{f}^l aux résidus améliore \hat{f} dans les zones où il ne fonctionne pas bien :

$$\hat{f}(x) \leftarrow \hat{f}(x) + \lambda \hat{f}^l(x)$$
(4.10)

Le paramètre de retrait λ contrôle la vitesse d'apprentissage d'un "arbre boost". En mettant à jour les résidus actuels :

$$r_i \leftarrow r_i - \lambda \hat{f}^l(x_i) \tag{4.11}$$

il est possible de donner l'"arbre boost" :

$$\hat{f}(x) = \sum_{l=1}^{L} \lambda \hat{f}^{l}(x)$$
(4.12)

AdaBoost Un exemple typique est AdaBoost : il s'agit d'une approche d'apprentissage automatique basée sur l'idée de créer une règle de classification très précise en combinant de nombreuses règles relativement faibles et inexactes [48] :

$$\hat{f}(x) = \begin{cases} 1 & si & \sum_{t=1}^{T} w_i^t h_t(x) \\ 0 & sinon \end{cases}$$
(4.13)

où h_t est le classifieur faible et w_i^t est le poids.

4.2.7 Avantages et inconvénients

L'absence d'une étape d'apprentissage fait de kNN un classifieur simple à être utilisé. Il n'exige aucune connaissance préalable sur la distribution des données, parce qu'il n'utilise que l'information de distance. Comme limitation du classifieur, on peut remarquer, qu'il est sensible au bruit. Il s'agit de la variation des échantillons d'apprentissage et de test [91].

L'une des limitations de SVM est le choix du noyau. Ce choix est à faire par l'expérimentateur et il est essentiel pour la bonne performance des SVM. Comme autre inconvénient on peut mentionner la dépendance de la taille des données pendant l'étape de l'apprentissage et la rapidité du classifieur [91].

Le plus grand inconvénient des kNN est la possibilité du modèle d'être trompée par des attributs non pertinents, sans application des pondérations aux données [66]. Par contre les arbres de décision sont capables de les ignorer.

Les arbres de décision sont aussi relativement rapides. L'inconvénient principal est qu'ils ont tendance à "surapprendre" les résultats. Généralement, ils ne présentent pas la précision de classification des autres approches de classification.



FIGURE 4.1: Résultats avec des classifieurs différents (issu de Christoff *et al.* [28]).

4.3 Comparaison des classifieurs selon différentes caractéristiques

Il est question de résoudre un problème de classification avec deux classes : "cratère" et "non cratère". Pour qu'on puisse trouver le classifieur le plus efficace pour apprentissage, on doit les mettre dans les mêmes conditions. On doit donner à chacun le même ensemble de données. Ensuite, on doit effectuer la classification pour un même ensemble de données de nouvelles régions et vérifier si les classes assignés correspondent.

Sur la Figure 4.1 est illustrée une analyse comparative de l'approche : celles de Di *et al. i.e.* une combinaison de trois caractéristiques ainsi que notre proposition. Cette représentation graphique des résultats est obtenue en utilisant différents classifieurs en fonction de quatre courbures principales et DEM. Les barres représentent la précision des classifieurs. Les lignes colorées représentent la précision de l'approche proposée pour chaque courbure et chaque DEM. On peut voir que le descripteur de fonctionnalité proposé est très stable avec de petites déviations. Les résultats en précision pour les différentes courbures sont très similaires à ceux de Di *et al.*, À une exception près - avec réseau de neurones, nous obtenons de meilleurs résultats et une précision de 98.8% sur k_2 .

4.4 Validation croisée

Généralement, il existe deux types d'erreurs : l'erreur d'apprentissage et l'erreur du test. Pour l'apprentissage, les N membres d'un ensemble de données M sont utilisés. Ensuite, le classifieur entraîné est utilisé comme ensemble initial de poids W_0 pour la validation. On divise aléatoirement l'ensemble des données en plusieurs sous-ensembles de taille approximativement égale : $M_j(j = 1, 2, ..., k)$, où k est le nombre des groupes. Le classifieur W_o est formé k fois avec chacun des ensembles de données [71]. On prend le premier sous-ensembles M_j pour la validation et le reste de k - 1servent à ajuster la méthode. Cette procédure se répète k fois et à chaque fois, l'ensemble de validation est différent. Pour chaque observation, l'erreur de test MSE_j est calculée. L'erreur de validation totale CV est calculée en faisant la moyenne de toutes les erreurs MSE_j , pour j = 1, 2, ..., k [66] :



FIGURE 4.2: L'apprentissage dans un perceptron multicouche. Le signal se propage vers l'avant et l'erreur en arrière (issu de Kasabov [71]).

$$CV_k = 1/k \sum_{j=1}^k MSE_j \tag{4.14}$$

L'objectif principal est de déterminer dans quelle mesure la procédure d'apprentissage estime l'erreur réelle du test MSE [66].

(

4.5 Notre choix de classifieur : le perceptron multicouche (MLP)

Le classifieur optimal choisi pour notre problématique est le réseau de neurones, parce qu'il donne une liberté avec le choix de nombre des neurones. Le réseau est flexible. Dans le cas général, il peut travailler avec n classes. Notre tâche consiste en une classification avec n = 2 : classe "cratère" et classe "non cratère". De plus, nous avons fait une étude expérimentale avec n = 3, pour effectuer une analyse des cratères détectés. Les cratères sont classifiés comme "valides", "dégradés" ou "secondaires".

4.5.1 Le perceptron multicouche

Pour la classification des échantillons de test, nous avons utilisé le réseau de neurones à propagation avant avec apprentissage supervisée, détaillé par Kasabov [71].

À chaque itération d'apprentissage, l'algorithme d'apprentissage est organisé en deux passages. Au cours de la passe directe, la direction des entrées se propage à travers les couches intermédiaires vers la couche de sortie. Durant la passe en arrière, l'erreur est calculée à chaque sortie et se propage en arrière pour calculer les modifications des poids. Cet algorithme est schématisé à la Figure 4.2. Lors du retour, une erreur E_{rri} pour un nœud intermédiaire *i* est calculée en multipliant les erreurs E_{rrj} de tous les neurones *j* auxquels le neurone *i* est relié par les poids correspondants w_{ij} . Pendant la passe en arrière, l'erreur E_{rri} est utilisée pour ajuster les poids des neurones *k* de la couche précédente, qui sont connectés au neurone *i*. Cette procédure d'apprentissage est répétée avec les mêmes exemples d'apprentissage jusqu'à ce que l'erreur globale *E* soit suffisamment faible [71].

4.5.2 Le processus d'apprentissage

Pendant le processus d'apprentissage, nous avons pris en compte les problèmes décrits par Kasabov [71] :

- L'apprentissage doit être objectif à partir d'ensemble de données. Si l'information contenue dans les données n'est pas suffisante, le système ne peut pas "apprendre" beaucoup.
- Le système apprend par des caractéristiques pertinentes pour trouver une solution, mais ne peut pas tout apprendre.
- Le test du processus d'apprentissage se fait généralement par la mesure d'erreur.

Les vecteurs caractéristiques, calculés dans la section 3.6 et les étiquettes de classe sont injectés dans le réseau de neurones en tant qu'entrées. L'idée du modèle proposé du réseau de neurones est d'être indépendant du nombre de vecteurs caractéristiques. Le nombre de neurones cachés dans la couche cachée doit être défini. Le nombre optimal de ces nœuds est une tâche difficile [71]. L'une des solutions que Kasabov propose et qui est utilisée dans ce travail est le réseau de neurones en croissance ("growing neural network"). Il s'agit d'un réseau de neurones qui démarre la procédure d'apprentissage avec un petit nombre de nœuds cachés. Le nombre peut augmenter, sous réserve de l'erreur calculée. Après plusieurs tests de valeurs pour le nombre de nœuds cachés, les meilleurs résultats ont été obtenus avec 50 neurones cachés.

La fonction d'activation est le troisième paramètre de configuration du réseau. Le gradient de propagation arrière conjugué ("conjugate back propagation gradient") est utilisé pour l'apprentissage du réseau, et offre le meilleur compromis entre une bonne classification et la performance selon Møller [99].

Nous avons estimé à quel point le modèle peut être entraîné par certaines données et ensuite s'il effectue une classification fiable en utilisant de nouvelles données fournies (c'est-à-dire des données avec lesquelles il n'a pas été entraîné). Pour s'assurer que les résultats de validation ne soient pas en sur-apprentissage ("overfitted") en raison de l'optimisation des données [61], la validation croisée k-fold ("k-fold cross validation") est utilisée pour ajuster les paramètres du modèle. Un ensemble de tests indépendants correspond à la division de la base de données en trois ensembles : apprentissage, validation et test [119].

Des valeurs de k égales à 5 ou 10 sont typiquement choisies [66, 71]. La base de données, obtenue dans la section 3.6 pour former et tester le réseau de neurones, contient 1505 échantillons positifs (qui contiennent un cratère) et un même nombre d'échantillons négatifs (qui ne contiennent pas de cratères). Les régions de la classe "non cratère" ne chevauchent pas les régions "cratère". Nous voulons fournir une sortie de k modèles ajustés qui sont moins corrélés les uns avec les autres. Avec cette hypothèse, nous avons choisi k = 5 car le chevauchement entre les ensembles d'apprentissage dans chaque modèle est plus petit. Ces 5 instances du modèle entraîné ne sont pas utilisées pour faire une classification réelle. Toutes les données doivent trouver le meilleur modèle possible. La distribution des trois ensembles est définie sur 80%, 10% et 10%. Quatre plis ou 80% sont utilisés pour la formation. À notre connaissance, il n'existe pas de résultats de recherche qui suggèrent que les pourcentages pour la validation (V) et les pourcentages

		Classe estimée		
	Classes	Cratère	Non cratère	
Classo róollo	Cratère	1372	133	
Classe leelle	ClassesCratèreNéelleCratère1372Non cratère29	1476		

TABLE 4.1: Résultats d'apprentissage du réseau de neurones.

pour le test (T) doivent être identiques ou différents. Nous avons utilisé la validation croisée imbriquée ("nested cross-validation"), où les 20 % sont également divisés en V = T = 10.

Une fois l'étape d'apprentissage terminé, les classes liées à chaque échantillon sont obtenues comme une sortie qui représente la détection (ou non) de cratères sur la *newArea* spécifique.

Ce réseau entraîné sera utilisé pour la détection des cratères dans le chapitre 5.

4.5.3 Résultats d'apprentissage

Le classifieur, décrit dans la section 4.5, a pour entrée les caractéristiques calculées de zones A, B et C (voir Figure 5.1). La moyenne des résultats du classification en validation croisée k = 5 est montrée dans le Tableau 4.1. Les régions étiquetées "cratère" et classifiées correctement par le réseau de neurones sont au nombre de 1372 sur un nombre total de 1505 échantillons; et pour la classe "non cratère" le nombre de correctement classifiés est 1476 sur un nombre total de 1505 échantillons. Les échantillons de la classe "cratère" considérés comme "non cratère" sont au nombre de 133. Actuellement, en plus de ces 1476 régions, le réseau devrait assigner 29 régions supplémentaire, mais elles sont classifiés comme ayant un cratère. Pour la précision globale du système, on obtient 91.1%. On doit souligner, que ces pourcentages diffèrent des résultats, montrés dans la section 4.3, parce qu'ils étaient obtenus avec une autre ensemble d'apprentissage, ayant de nombre d'échantillons réduits.

4.6 Conclusion

Dans ce chapitre, nous avons présenté certaines méthodes d'apprentissage automatique et de classification supervisée. Nous avons comparé les performances des techniques présentées avec des caractéristiques extraites du chapitre 3. Nous avons montré que notre approche générique produit un classifieur stable et, selon le classifieur choisi, de très bons résultats par rapport aux autres méthodes issues de la littérature. Cela peut sans doute être utilisé pour détecter des formes rondes plus génériques sur des maillages et dans d'autres contextes.

Chapitre 5

Extraction et analyse des cratères détectés

Sommaire

5.1	Détec	tion des régions des cratères	59
5.2	Extra	ction des bords des cratères	61
5.3	Const	truction du catalogue Christoff et al.	62
	5.3.1	Correspondance des cratères	62
	5.3.2	Effacement des cratères similaires	63
5.4	Analy	yse de catalogue Christoff <i>et al.</i>	64
	5.4.1	Analyse selon le diamètre	64
	5.4.2	Analyse selon la position	65
	5.4.3	Analyse selon le type de cratère	65
5.5	Conc	lusion	68

D^{ANS} ce chapitre, nous proposons de tester la performance de notre méthode. Nous décrivons la manière de procéder à l'extraction des bords des cratères et nous effectuons une analyse des cratères détectés. Ce chapitre est basé sur nos publications [29, 30].

5.1 Détection des régions des cratères

Après avoir entraîné le classifieur, nous le testons sur de nouvelles surfaces de Mars pour sélectionner ces régions, qui contiennent des cratères.

La Figure 5.1 montre les régions que l'on utilise, entourées par des rectangles gris. Les zones A, B et C sont les régions à partir desquelles l'ensemble d'apprentissage est construit (des explications détaillées sont données dans le Chapitre 4, section 4.5) et les zones 1, 2 et 3 constituent les régions de validation pour l'approche proposée. C et 1 sont une seule et même région, utilisée à la fois pour l'apprentissage et pour la validation, à cause des comparaisons décrites dans la subsection 6.2.3 . Dans le Tableau 5.1, une information supplémentaire sur les coordonnées de toutes les zones est donnée.

La région 1 est située autour du cratère Herschel, la région 2 est située autour du cratère Huygens et la région 3 est située au sud de Valles Marineris. Tous les sites sont exactement les mêmes que ces utilisés par la méthode de Di *et al.* [41]. De plus, la dalle 1 est aussi prise pour validation par Bue et Stepinski [17]. Cela va nous permettre d'effectuer une comparaison avec ces deux méthodes. Il est important de noter que les sites 1 et 2 ont une distribution similaire de cratères. Ces régions sont difficiles à traiter,



FIGURE 5.1: Carte de Mars : les zones d'apprentissage (A, B et C) et les zones de validation (1, 2 et 3).

Dalle	Haut	Bas	Gauche	Droit	Surface (km^2)
А	0.00°	$13.00^{\circ}N$	$25.00^{\circ}W$	$0.01^{\circ}W$	1, 142, 158
В	$13.00^{\circ}N$	$26.00^{\circ}N$	$25.00^{\circ}W$	$0.01^{\circ}W$	1, 142, 158
1	$7.57^{\circ}S$	$18.42^{\circ}S$	$114^{\circ}E$	$142^{\circ}E$	1,067,464
2	$8.05^{\circ}S$	$18.89^\circ S$	$37.69^{\circ}E$	$65.69^{\circ}E$	1,064,565
3	$14.85^\circ S$	$25.69^\circ S$	$88.87^{\circ}W$	$61.26^{\circ}W$	1,050,006

TABLE 5.1: Informations supplémentaires sur les dalles.

contrairement à la région 3, parce que les cratères sont densément répartis, ils se chevauchent et ils sont fortement dégradés.

Pour la détection des cratères de tailles différentes, les courbures $k_{2,G}$ sont redimensionnées et une pyramide est construite. La taille de chaque couche est réduite avec 1.2^{-1} fois la taille de la couche précédente. La première couche de la pyramide représente les dimensions originales du k_{2,G} et la dernière est la plus petite région de cratère avec une largeur et une hauteur d'au moins 20 pixels. Sur chaque couche, une fenêtre de 20×20 pixels est déplacée à chaque fois de 3 pixels pour balayer la courbure minimale $k_{2,G}$. Pour chaque fenêtre, les caractéristiques sont extraites en utilisant l'algorithme décrit dans la section 3.6. Les valeurs locales Max, Mean et *Min* sont utilisés pour marquer les domaines définis dans le Tableau 3.1. Chaque fenêtre indique une nouvelle entrée pour le réseau de neurones. Il a pour but d'estimer la probabilité pour chaque entrée d'appartenir à l'une des deux classes "cratère" ou "non-cratère". La décision pour la région est prise sur la base de la classe ayant la plus grande probabilité. Il s'agit des régions d'intérêt pour lesquelles on va chercher à extraire les bords des cratères.

5.2 Extraction des bords des cratères

En cas de détection d'une région qui contient des cratères, les coordonnées de la zone newArea dans laquelle un cratère est détecté sont sauvegardées. Ensuite, la transformée de Hough circulaire (CHT) [20] est appliquée sur les images $k_{2,G}$ correspondantes à ces coordonnées enregistrées. Il s'agit d'une technique d'extraction de caractéristiques spécifiques, dont le but est de dessiner des cercles probables à partir d'entrées d'images imparfaites. Les candidats pour tracer des cercles sont produits par des *votes* dans l'espace des paramètres de Hough, puis les maxima locaux sont sélectionnés dans une matrice dite d'accumulateurs.

Sur les régions où des cratères sont détectés, Di *et al.* [41] Bue et Stepinski [17] ont proposé de positionner un masque sur chaque pixel du DEM original. Pour chaque masque, la moyenne locale est calculée, ainsi que les plus grands écarts de cette moyenne au minimum ou au maximum des valeurs. Ces plages de valeurs, affectées à chaque région, sont utilisées pour calculer le seuil de binarisation de l'image, et on peut s'attendre par conséquent à un changement de topologie sur la région. Di *et al.* [41] et Bue et Stepinski [17] effectuent ensuite une fermeture morphologique, suivie d'un éclaircissement afin de réduire les lignes de l'image à la largeur d'un pixel. L'image transformée est prête à subir la CHT.

Délimitation du bord du cratère : dans ce travail de recherche, pour économiser du temps et de la mémoire, les paramètres sont directement indiqués, en fonction du contraste, pour adapter le cercle (la zone sombre de l'image). Le rayon est compris dans l'intervalle [5, 10], en supposant que l'on obtient un cercle caractérisant un cratère dans la sous-fenêtre 20×20 . Les cercles qui sortent de la fenêtre sont exclus (les coordonnées de centre sont dans la fenêtre 20×20 , mais le bord est en dehors). Les points noirs locaux de l'accumulateur tridimensionnel fournissent la position et le rayon du cratère. L'exhaustivité du cercle est donnée par l'accumulateur. La réduction de la condition de seuil génère plus de cercles faux positifs, donc la valeur de seuil minimale est fixée à 0.1. Les données du cercle ainsi obtenues (coordonnées du centre estimé et diamètre) sont finalement réajustées et redimensionnées en fonction des coordonnées de la carte.

Sur la Figure 5.2 a), les résultats obtenus de la transformée de Hough circulaire appliquée sur l'image $k_{2,G}$, contenant un cratère, sont montrés. Le cercle bleu est le bord du cratère estimé et la ligne rouge est le rayon du cercle r, estimé à 14.8 km.

Délimitation précise du bord du cratère : dans la nature, les diamètres des cratères ne sont pas des puissance de 2 et le redimensionnement ne donnera pas de bons centres et ni les limites exactes. À partir de la liste des rayons et des centres des cratères ré-échantillonnés, la CHT est à nouveau appliquée. On cherche à obtenir un rayon dans les contraintes [r*0.5, r*1.5] et on change le seuil minimal à 0.001.

Les résultats de cette délimitation sont présentés sur la Figure 5.2 b). On obtient le bord du même cratère à 13.06 km, visualisé par le cercle vert. Ceci ne modifie presque pas le centre du cratère.



FIGURE 5.2: Résultat après CHT sur l'image $k_{2,G}$: a) bord ré-échantillonné, b) bord délimité.



FIGURE 5.3: Cratères «similaires» selon leur zone de chevauchement relatif. Le cercle rouge est "Crater 1" et le cercle vert "Crater 2" (issu de Christoff *et al.* [30]).

5.3 Construction du catalogue Christoff et al.

Pendant la CHT, en raison des multiples couches existantes et de la fenêtre glissante, un même cratère peut être assigné avec des coordonnées de centre et/ou un rayon légèrement différents. Ces répétitions sont détectées et supprimées.

5.3.1 Correspondance des cratères

Salamunićcar et Lončarić proposent un processus de correspondance automatisé [124] pour déterminer si deux entrées de la liste des cratères obtenus sont identiques ou différentes. Ils proposent un processus relatif de correspondance pour deux cratères : "Crater 1" et "Crater 2". À partir de la liste des cratères stockés, où X, Y sont les coordonnées du centre du cratère et R est le rayon, deux cratères sont considérés comme appartenant au même cratère s'ils satisfont l'Équation 5.1 et l'Équation 5.2 :

$$f_{diff} = max(\frac{r_1}{r_2} - 1, \frac{d}{r_2})$$
(5.1)

$$f_{diff} < f_c \tag{5.2}$$

Salamunićcar et Lončarić introduisent la définition de mesure de la différence de cratère f_{diff} [124]. Les r_1 et r_2 sont respectivement les rayons de "Crater 1" et de "Crater 2", où $r_1 \ge r_2$. La distance entre les centres du cratère est marquée par d et $f_c \ge 0$ est le facteur de différence des cratères. Sur la Figure 5.3, trois cas particuliers de f_{diff} sont montrés : $f_{diff} = 0.25$ (a), $f_{diff} = 0.50$ (b), $f_{diff} = 1.00$ (c). Le cas (c) satisfait les dépendances suivantes :

$$f_{diff} = \frac{r_1}{r_2} - 1 = \frac{d}{r_2} \tag{5.3}$$

Le seuil d'une correspondance correcte f_c peut être défini après une vérification de toutes les combinaisons possibles dans le catalogue. Le plus petit f_{diff} , trouvé entre différents cratères, définit f_c qui doit être encore plus petit. Ce n'est que lorsque le maximum est utilisé dans l'Équation 5.1 que la contribution de la différence entre le rayon et la contribution de la distance entre les coordonnées peut aller jusqu'à la f_c . Cela maximise la probabilité de correspondance correcte et fait que l'Équation 5.1 est optimale.

Burl *et al.* [18] et Wetzler *et al.* [150] utilisent une mesure similaire. La zone relative de chevauchement mesure le rapport entre la zone d'intersection entre les deux cercles et la zone du plus grand des deux. Trois cas particuliers sont à considérer :

- Un cratère plus petit est complètement à l'intérieur d'un cratère plus grand : les deux mesures sont calculées en utilisant seulement la différence entre leurs rayons. Dans ce cas, en partant d'une mesure, il est possible d'en calculer une autre, et vice-versa, afin qu'elles puissent être considérées comme équivalentes.
- Un cratère plus petit est au moins en partie à l'extérieur d'un cratère plus grand : les mesures sont différentes.
- Un cratère est complètement à l'extérieur d'un autre, ils ne se chevauchent pas : la mesure est seulement définie par l'Équation 5.1.

L'Équation 5.1 est importante lors du processus d'enregistrement du catalogue [124] afin de s'assurer que certains cratères ne sont pas involontairement manqués (comme cela est montré sur les Figures 2.3 a) et b) : les plus petits cratères des catalogues de Barlow et Rodionova se diffèrent complètement à la topographie sous-jacente et dans certains cas pour plusieurs cercles).

5.3.2 Effacement des cratères similaires

On applique les critères de similarité deux fois : une fois après chaque couche, qui sert pour la construction de la liste créée dans la Section 5.2; la seconde est pour chaque élément de cette liste. On efface les répétitions, en utilisant les équations suivantes :

$$x_{cr} = \frac{\sum_{1}^{m} x_{m}}{m}; \qquad y_{cr} = \frac{\sum_{1}^{m} y_{m}}{m}; \qquad r_{cr} = \frac{\sum_{1}^{m} r_{m}}{m}$$
(5.4)

En cas de coïncidence, les coordonnées (x_{cr}, y_{cr}) et le rayon r_{cr} pour un cratère sont calculés à partir de la division de la somme des coordonnées particulières et le rayon par le nombre d'entrées similaires m.

5.4 Analyse de catalogue Christoff *et al.*

Sur la Figure 5.4 est présentée une petite région de Mars avec les résultats obtenus par notre méthode. Les cercles blancs montrent les positions des bords des cratères trouvés et leurs tailles.



FIGURE 5.4: Les cratères précis avec leurs rayons.

5.4.1 Analyse selon le diamètre

Pour la validation de l'efficacité de l'approche proposée dans cette thèse, l'identification précise du diamètre du cratère d'impact sera un critère de très grande importance.

Dans le Tableau 5.2, une étude statistique des diamètres des bords est effectuée, avant et après la délimitation précise. Ils sont rangés dans trois groupes en fonction de leur taille. Dans le premier entrent tous les cratères ayant un diamètre $D \le 6$ km. Ils correspondent aux petits cratères. La seconde catégorie correspond aux cratères moyens, ayant un bord en diamètre 6 < D < 25 km. Dans la troisième catégorie sont rangés les grands cratères ($D \ge 25$ km). Il est à mentionner que le nombre total après la seconde CHT diminue pour les trois dalles (8%, 9%, 13% respectivement).

En ce qui concerne la répartition des diamètres, les petits cratères (entre 59% - 67%) sont les plus nombreux. Après la seconde CHT, le nombre des diamètres plus petit ou égal à 6km diminue légèrement, mais le pourcentage augmente (64 - 77%). En ce qui concerne les grands cratères, ils représentent le plus petit nombre de détéctions (2-3%). De même que pour les petits cratères, la quantité change légèrment, mais la population diminue jusqu'à 1-3%. Les changements majeurs concernent les cratères ayant un diamètre entre 6 et 25 km. Leur nombre passe à 2% pour la dalle 1, 15% pour dalle 2 et 9% pour dalle 3.

avant					â	après		
Dalla	Nombre	D < 6 km	D > 6km,	D > 25 km	Nombre	D < 6 km	D > 6km,	D > 25 hm
Duile	total	$D \leq 0 \kappa m$	D < 25 km	$D \ge 25 \kappa m$	total	$D \leq 0 \kappa m$	D < 25 km	$D \ge 25 \kappa m$
1	1142	704	404	34	1052	674	345	33
2	1844	1080	728	36	1670	1074	561	35
3	982	662	303	17	858	657	188	13

TABLE 5.2: Statistiques des cratères détectés automatique-ment sur les sites 1, 2, 3, avant et après la délimitation pré-cise du bord du cratère.

5.4.2 Analyse selon la position

On peut considérer les types de position de cratères suivants :

- Selon la position sur la dalle : notre approche est capable de détecter les cratères qui sont entièrement dans la région examinée.
- Selon le chevauchement : la distribution des cratères influe sur la détection pour les zones ayant une grande distribution des cratères de tailles différentes.
- La position des centres coïncide avec la topographie de la surface planétaire (voir Figure 5.4).
- Il est possible que deux cratères différents soient fusionnés s'ils ont des tailles similaires et se trouvent très proches l'un à l'autre [124].

5.4.3 Analyse selon le type de cratère

Le nombre exact de cratères d'impact doit être pris en compte afin de dater la surface. Cependant, des cratères secondaires et dégradés ont affecté les statistiques de dénombrement des cratères de la surface.

Il n'existe pas à notre connaissance de système commun et standardisé qui regroupe les cratères en fonction de leurs caractéristiques morphologiques. Pour les cratères lunaires, il y a peu de méthodes d'identification des cratères d'impact selon leur structure morphologique. Pour les simples cratères lunaires, Vijayan *et al.* [144] détectent, effectuent la classification et extraient des informations contextuelles. La classification des cratères est effectuée selon le type du fond (rond ou plat). Les auteurs indiquent la présence d'éjecta et l'associent avec le cratère correspondant. Ils évaluent également l'état de dégradation du cratère. Toujours pour les cratères d'impact lunaires, l'étude de Chen *et al.* [24] propose un système d'index descriptif basé sur les caractéristiques morphologiques et un standard de description structurée.

Nous effectuons l'analyse les cratères détectés par l'algorithme proposé dans ce manuscrit. Le résultat de classification morphologique automatique catégorise les cratères d'impact dans trois catégories. Un expert a assigné une morphologie de cratère d'impact à une classe donnée, après inspection visuelle [81] :

 Cratères valides : cette classe inclut des cratères frais et des structures d'impact qui ne sont pas suffisamment modifiées à cause de l'érosion ou soumis à d'autres influences climatiques. Elle inclut des cratères simples (voir Figure 1.3) et des cratères complexes (voir Figure 1.4). Les cratères dans cette catégorie peuvent contenir ou pas des couvertures d'éjecta. Il est possible que cette catégorie contienne



FIGURE 5.5: Les types des cratères (issu de Christoff *et al.* [29]).

des cratères secondaires, s'ils ne sont pas associés à un impact primaire ou à un groupement. Un échantillon de cratère qui appartient de cette classe est montré sur la Figure 5.5 a1). Cette catégorie «valide» contient des cratères d'impact, qui doivent être inclus dans le nombre de cratères pour une datation de la surface.

- 2. Cratères dégradés : les cratères ici présents ont un bord partiellement ou complètement détruit, ou bien leur intérieur est rempli de matériaux sédimentaires [7]. L'élévation de leur fond est (presque) au même niveau que la surface d'origine (la surface avant l'impact). Un exemple de tel cratère dégradé est présenté sur la Figure 5.5 b1). Ils doivent être ignorés du nombre de cratères, qui vont estimer l'âge, parce que la formation de ces cratères est apparue avant la surface datée [11].
- 3. Cratères secondaires : cette catégorie contient des structures d'impact, associés à un impact primaire (par exemple des rayons aux cratères primaires environnants) ou à un groupement (Figure 1.6) [121]. Le diamètre des cratères secondaires, habituellement, ne dépasse pas 1 km [95]. Cela rend sa reconnaissance très difficile et ce n'est qu'après une analyse régionale qu'ils peuvent être distingués. La Figure 5.5 c1) illustre un exemple d'échantillon de cette catégorie. Dans leurs recherches, Hartmann et Neukum [60] mentionnent que les cratères secondaires ont affecté les statistiques du nombre de cratères, qui sont utilisées pour faire la datation de la surface. Les cratères secondaires sont d'origine non primaire, ce qui signifie qu'ils doivent être omis du comptage de cratères.

La Figure 5.5 montre les trois types de cratères que l'on catégorise. La lettre "a" est associée au cratère "valide", "b" est associée au cratère dégradé et "c" est associée au cratère secondaire. Pour chaque classe, le label "1" indique le maillage 3D, les courbures minimales calculées sont indiquées par



FIGURE 5.6: Les résultats de catégorisation (issu de Christoff *et al.* [29]).

le label "2", et le label "3" indique les résultats de la méthode de quantification locale. Un total de 100 échantillons a été considéré dans chaque classe.

On ne dispose pas d'assez d'échantillons pour avoir une base de données ayant le même nombre d'objets dans chacune des trois classes. Pour cette raison, nous utilisons un "Residual Network" avec 50 couches (ResNet50) [62].

Une fenêtre glissante traverse l'échantillon d'entrée avec une convolution appliquée à chaque pas. Chaque pixel est multiplié par la valeur de filtre correspondante. L'image est formée à partir des résultats de ces convolutions.

Les sorties des couches précédentes (des vecteurs x) sont utilisées comme des entrées pour les couches entièrement connectées, qui sont construits de nombreuses unités. Il s'agit d'une transformation $h_{\theta}(x)$ dont le résultat est une valeur unique où θ est le vecteur de poids. On utilise une unité linéaire rectifiée (ReLU) : $h_{\theta}(x) = max(0, \sum_{i=0} \theta_i x_i)$ [32] comme fonction d'activation, ce qui accélère beaucoup la convergence du gradient.

On construit une architecture comme celle proposée par Zagoruyko et Komodakis[152]. Elle est une variante de l'architecture originale de He *et al.* [62]. Chaque entrée représente une image 20×20 , qui est traitée avec trois couches convolutives avec 8, 16 et 32 filtres, chacun avec une taille 3×3 pixels. Chaque filtre passe sur l'image précédemment filtrée dans une fenêtre glissante avec un pas de 1 pixel. Le bloc résiduel est complété après chaque convolution par la normalisation des lots [65] et la fonction d'activation (ReLU). On utilise une couche « mise en commun » ("de pooling"), qui sous-échantillonne l'image). Chaque élément dans la couche entièrement connectée représente le résultat de $h_{\theta}(x)$. On utilise une régression exponentielle normalisée (softmax) pour obtenir la distribution de probabilité entre les classes [14]. La classification donne trois sorties possibles, qui correspondent à l'une des trois classes.

On utilise la validation croisée, qui divise la base de données originale en k = 5 échantillons . La base de données est divisée comme suit : 80% de chaque classe ont été choisis au hasard comme échantillons d'apprentissage et le reste (20%) ont été utilisés pour le test. Les résultats de la classification sont présentés dans le Tableau 5.3 et sur la Figure 5.6.

La base de données et les résultats de cette classification sont montrés

		Valides	Dégradés	Secondaires	
	Valides	18	0	2	90%
Classe	Dégradés	2	17	1	85%
reelle	Secondaires	0	2	18	90%
		90%	89.5%	85.7%	88.3%

TABLE 5.3: La matrice de confusion des résultats obtenus.

sur la Figure 5.6. Les échantillons d'apprentissage sont marqués en rectangles jaunes. Il représent 80% de la base de données. Le résultat du test (les 20% restants) sont caractérisés comme suit : 18 cratères, montrés avec des rectangles verts, ont été classés comme «valides». Ils correspondent à une catégorisation correcte dans classe «valides» de 90% (voir le Tableau 5.3). Le même nombre de 18 (et pourcentages) ont été catégorisés comme cratères «secondaires» de troisième classe (rectangles bleus sur la Figure 5.6. Le taux de reconnaissance de la classe «dégradé» et 85% (ou 17 cratères) (voir Tableau 5.3). Sur la Figure 5.6, ils sont illustrés par des rectangles blancs. Les cratères qui sont en rouge sont mal classées (voir la Figure 5.6). A chaque classe sont attribuées une couleur et une chiffre : «valide» - vert - 1; «dégradé» - blanc -2 et «secondaire» - bleu - 3. A côté de chaque rectangle rouge, le chiffre indiqué correspond à la classe réelle et la couleur exprime la catégorie prédite (voir Figure 5.6) en fonction des classes prédites de la matrice de confusion (voir Tableau 5.3). La précision du notre système est de 88.3%. Nous avons essayé de faire la catégorisation avec le réseau de neurones artificiels, utilisé pour la détection. Comme résultat, la précision était de 58.2%. Ces résultats obtenus ont montré que le réseau ne distingue pas les cratères «valides» des cratères «secondaires» en raison de leurs structures morphologiques similaires.

5.5 Conclusion

Ce chapitre était consacré à l'extraction des bords des cratères par l'application de la CHT sur les régions de cratère détectés, obtenus après la classification des régions, en utilisant le classifieur décrit dans le chapitre 4. Le catalogue généré indique la position et la taille des cratères détectés. Nous avons fait une analyse des résultats obtenus. Les centres sont bien positionnés par rapport à la topographie. Les cercles suivent les contours des bords. Avec notre approche, nous obtenons un grand nombre de détections des cratères de petite taille (plus petit de 6 km). La validation et l'évaluation de notre catalogue sont données dans le chapitre 6.

En supplément, nous avons fait une étude sur la morphologie des cratères détectés, en essayant de les catégoriser dans trois classes. À notre connaissance, il n'y a pas d'autres travaux qui effectuent une telle classification. Cela ne nous permet donc pas de comparer nos résultats avec d'autres méthodes existantes.

Chapitre 6

Résultats et validation

Sommaire

6.1	Précision et évaluation des algorithmes de détection de						
	cratères	69					
6.2	Évaluation	70					
	6.2.1 Comparaison avec les autres catalogues	70					
	6.2.2 Facteurs de qualité	72					
	6.2.3 Comparaison avec d'autres méthodes	74					
6.3	Conclusion	77					

C E chapitre présente les résultats de la méthode de détection et caractérisation d'attributs géométriques sur les corps rocheux du système solaire. Les résultats ici présents s'appuient sur la publication [30], et nous avons intégré dans ce chapitre des résultats supplémentaires et une analyse plus approfondie.

6.1 Précision et évaluation des algorithmes de détection de cratères

Dans le cas idéal (un ADC parfaitement précis et à l'échelle invariable), le système de détection ne fournit aucune erreur sur l'efficacité de détection, la mesure, ou les deux. Ces erreurs peuvent influer sur le résultat final d'estimation de l'âge. Ils peuvent également affecter la détermination des caractéristiques morphologiques des cratères.

Les facteurs de qualité sont : le pourcentage de détection (voir Équation 6.1), le pourcentage de qualité (voir Équation 6.2) et le facteur de branchement ("branching factor") (voir Équation 6.3) développés par Shufelt [131] sont utilisés dans beaucoup d'autres ADC pour évaluer la performance de la méthode [5, 41]. La classification des cratères peut être : vrai positif (TP) qui représente le nombre de cratères réellement détectés, les faux positifs (FP) qui est le nombre des cratères détectés qui ne sont pas réellement des cratères et les faux négatifs (FN) représente le nombre de cratères réels non détectés.

Pourcentage de détection =
$$100 \times \left(\frac{TP}{TP + FN}\right)$$
 (6.1)

Pourcentage de qualité =
$$100 \times \left(\frac{TP}{TP + FP + FN}\right)$$
 (6.2)

$$Facteur \ de \ branchement = \left(\frac{FP}{TP}\right) \tag{6.3}$$

Pour évaluer ces algorithmes, de nombreux chercheurs se comparent à un catalogue de cratères identifiés manuellement (le plus souvent celui de Barlow ou de Robbins). La précision et la réponse à la variation de certains paramètres doivent être dans certaines limités (on a vu préalablement les décalages des coordonnées dans certains catalogues). Ces variables sont par exemple la taille des cratère, les unités géomorphologiques, ou le type de données d'élévation.

Il est important de mentionner que les différents auteurs utilisent différentes zones d'essais, différents catalogues, voire différentes mesures statistiques d'évaluation. Des différences dans les taux de détection peuvent aussi varier en fonction du type de données utilisé. Il n'existe pas encore un modèle homogène pour l'évaluation objective de ces systèmes [124].

6.2 Évaluation

6.2.1 Comparaison avec les autres catalogues

Nous allons comparer nos résultats avec ceux obtenus en utilisant deux méthodes : celle de Di *et al.* [41] et celle de Bue et Stepinski [17], qui utilisent deux catalogues différents comme "vérités-terrain" (ceux de Robbins [120] et Barlow [9] respectivement). Pour cette raison, dans cette section, nous allons examiner les cratères qui sont présents dans les deux catalogues et qui sont situés dans chacune des trois dalles 1, 2 et 3. De plus, on effectue une évaluation de nos cratères avec ceux qu'ils détectent.

Dans les Tableaux 6.1 et 6.2, une statistique est effectuée pour des diamètres des cratères dont les centres entrent des trois régions (1, 2 et 3). Ils sont divisés en trois groupes : petits, moyens et grands. Il faut noter que le nombre total ne correspond pas au nombre de tous les cratères qui sont situés dans les limites des dalles. On donne le nombre des cratères, qui ont un diamètre plus grand ou égal au plus petit diamètre détecté (*min*) par notre approche. Le diamètre *min* est différent pour chaque dalle (1.5 - 2km). Les petits cratères sont compris entre *min* $\leq D \leq 6$. Les cratères moyens ont un diamètre de bord 6 < D < 25 km et les grands cratères correspondent à $D \geq 25$ km.

Catalogue de Robbins Le catalogue de Robbins contient un nombre des cratères, montré dans la colonne 2 du Tableau 6.1. De même que pour notre catalogue, le nombre de petits cratères contient la majorité des diamètres (64% pour le site 2 et 59% pour le site 3). Le pourcentage des diamètres plus petit ou égal à 6 km est un peu plus petit (43%) pour le site 1 en comparant avec 6 < D < 25km (47%). Les grands cratères constituent 5 à 10% de tous les cratères dans les zones.

Dalle	Nombre $total \ge min$	$D \leq 6km$	$\begin{array}{l} D>6km,\\ D<25km \end{array}$	$D \ge 25 km$
1	796	340	378	78
2	1042	665	302	75
3	296	176	106	14

TABLE 6.1: Statistiques des cratères qui figurent dans le ca-
talogue de Robbins [120].

Dalle	Nombre $total \ge min$	$D \leq 6km$	$\begin{array}{l} D>6km,\\ D<25km \end{array}$	$D \ge 25 km$
1	499	111	318	70
2	432	88	275	69
3	202	78	113	11

TABLE 6.2: Statistiques des cratères qui figurent dans le catalogue de Barlow [9].

Catalogue de Barlow Le catalogue de Barlow est constitué principalement de cratères de taille moyenne (56 - 64%). On trouve ensuite les petits cratères (20 - 39%). Le pourcentage des grands cratères est toujours le plus petit (5 - 16%).

La Figure 6.1 illustre la même zone que celle de la Figure 5.4. Les bords des cratères, ici montrés en blanc, sont colorés en vert (ils figurent dans au mois un catalogue) et en rouge (ils ne sont pas considérés comme un cratère par l'expert). À gauche, les entrées qui figurent dans le catalogue de Robbins sont visualisées en mauve. On peut voir qu'il n'y a pas de grandes différences par rapport à la position des cercles. Regardons maintenant à droite, où les cratères notés dans le catalogue de Barlow sont montrés par des cercles jaunes. On peut observer qu'ils ne sont pas centrés. Nous alignons les coordonnées des centres avec une distance et une direction telle qu'elles sont représentées sur la Figure 6.1 (à droite) (avec des flèches bleues). La taille de nos diamètres varie entre 1 et 1.3 fois les diamètres des bords du cratère, définis par Barlow.

La Figure 6.2 montre un graphique du nombre de cratères cumulés en fonction du diamètre. Les diamètres qui figurent dans notre catalogue sont tracés par une courbe bleue. En mauve sont montrés les cratères du catalogue de Robbins et ceux de catalogue de Barlow sont visualisés en jaune.

L'identification précise du diamètre du cratère d'impact est un critère important pour valider l'efficacité de l'approche de détection. Les diamètres des cratères d'impact détectés automatiquement (1052 pour la dalle 1, 1670 pour la dalle 2 et 858 pour la dalle 3) et les catalogues préparés manuellement sur les régions de tests sont visualisés (voir la Figure 6.2). À partir des graphiques montrés, on peut voir que la courbe qui présente notre catalogue (en bleu) a les mêmes propriétés que celles des catalogues de Robbins et Barlow.

Les performances de notre méthode dans les trois régions d'essai diffèrent. Le site 1 et le site 2 sont très difficiles, en raison de la grande variété de taille de cratères, de géométries qui se chevauchent et d'une énorme dégradation. Cela affecte la détection de ces cratères qui ont un diamètre supérieur à $50 \ km$.



FIGURE 6.1: Comparaison de nos cratères avec ceux du catalogue de Robbins (à gauche) et du catalogue de Barlow (à droite).

	Cratères dans le catalogue de Barlow	Cratères trouvés par notre approche	Correctement détecté (TP)	Cratères appariés	Cratères omis (FN)	Faux positifs (FP)
Site 1	499	1142	987	336	163	155
Site 2	432	1844	1434	311	121	410
Site 3	202	982	708	188	14	274

TABLE 6.3: Évaluation quantitative standard de la qualité des cratères d'impact avec le catalogue de Barlow.

6.2.2 Facteurs de qualité

On a déjà montré que la délimitation précise détecte correctement les bords des cratères, mais le nombre des entrées diminue. On va explorer la qualité des deux listes pour prouver que la délimitation augmente aussi l'aptitude des détéctions.

Pour calculer les facteurs de qualité et évaluer la performance de notre méthode, on catégorise chaque cratère comme vrai positif (TP) ou faux positif (FP). Le nombre des TP est obtenu en utilisant le critère de similarité expliqué dans la section 5.3.1. Il est appliqué ici pour faire correspondre un cratère détecté à un cratère du catalogue de Barlow et du catalogue de Robbins. Pour valider tous les cratères et gagner du temps, nous avons apparié nos cratères avec deux autres catalogues : Rodionova [123] et Salamuniccar [127]. Le nombre des TP est complété par le nombre de cratères qui sont vérifiés manuellement et signés comme des cratères. Les entrées qui ne figuraient dans aucun catalogue et ne sont pas libérées comme des cratères

	Cratères dans le catalogue de Robbins	Cratères trouvés par notre approche	Correctement détecté (TP)	Cratères appariés	Cratères omis (FN)	Faux positifs (FP)
Site 1	796	1052	970	495	301	82
Site 2	1042	1670	1420	719	323	250
Site 3	296	858	696	268	28	162

TABLE 6.4: Évaluation quantitative standard de la qualité des cratères d'impact avec le catalogue de Robbins.



FIGURE 6.2: Différences des diamètres des cratères catalogués / détectés, respectivement pour a) site 1, b) site 2 et c) site 3. Les diamètres des bords des cratères (km) qui sont trouvés par notre approche sont tracés en bleu, les lignes jaunes sont ceux issus du catalogue de Barlow, et en mauve ceux du catalogue de Robbins.

constituent le nombre des faux positifs. Pendant le processus de similarité, pour chacun des deux catalogues, on compte le nombre de ces cratères qui ne s'apparient avec aucun de nos cratères. Cette somme indique la valeur de FN.

Les résultats de la détection avant la délimitation précise sont montrés sur la Figure 6.3 et le Tableau 6.3 [30]. Ils sont validés par le catalogue de Barlow. On doit tenir compte que tous les centres de cratères présentés dans le catalogue Barlow sont comptés, de même que ceux qui ne sont pas entièrement situés dans les limites de chaque zone. Les facteurs de qualité pour les cratères délimités sont montrés dans le Tableau 6.4 et la Figure 6.4. Nous avons décidé de faire la validation avec le catalogue de Robbins, pour que la comparaison avec la méthode de Di *et. al* soit plus juste. Dans les statistiques données, on exclut ces cratères, qui ne sont pas entièrement dans les régions traitées.

Sur les Figures 6.3 et 6.4, on visualise les cratères correctement détectés, qui correspondent au moins dans l'un des quatre catalogues à des cercles verts. Les cratères qui sont TP mais dont on n'a pas trouvé une similarité sont étiquetés manuellement comme étant des cratères et sont visualisé par des cercles bleus. Les cercles rouges représentent les faux positifs. Ces cratères qui sont appariés avec le catalogue de Barlow (sur la Figure 6.3) et avec le catalogue de Robbins (sur la Figure 6.4) sont rangés dans des rectangles verts. Les rectangles rouges entourent les cratères qui sont omis par notre méthode.

Les statistiques obtenues dans les Tableaux 6.3, 6.4 nous permettent de calculer les facteurs de qualité pour chaque région. Le pourcentage de détection, le pourcentage de qualité et le facteur de branchement sont obtenus par les Équations 6.1, 6.2 et 6.3. Les résultats affichés dans le Tableau 6.5 sont obtenus à partir d'information dans le Tableau 6.4. Les résultats pour la comparaison avec l'approche de Bue sont estimé avec le Tableau 6.3.

6.2.3 Comparaison avec d'autres méthodes

Les performances de notre méthode sont diminuées par la complexité des régions 1 et 2, ce qui est aussi le cas pour les autres ADC. Les résultats montrent (cas de la région 3) que l'on a de meilleurs facteurs de qualité si les cratères dans la région sont bien séparés, moins recouverts et moins dégradés.

Le tableau 6.5 est un comparatif entre notre approche et celle décrite dans Di *et al.* [41]. Pour les trois régions, nous avons un taux de détection légèrement supérieur. Nous obtenons un meilleur facteur de ramification ("branching factor", le plus petit étant meilleur) et la qualité des cratères trouvés est améliorée. Il convient de noter que : 1) les résultats de Di *et al.* [41] sont obtenus en utilisant le catalogue de Robbins [120]; 2) pour détecter les régions ayant des cratères, ils utilisent une fenêtre glissante sur les couches de données du type DEM, en commençant par la taille originale de DEM. Chaque nouvelle couche est redimensionnée 1.2^{-1} fois; 3) ils utilisent aussi deux régions supplémentaires d'où ils prennent des échantillons positifs et négatifs pour entraîner le classifieur (qui sont différents de ceux qu'on a utilisés). Après cela, ils testent le classifieur déjà entraîné, sur les régions 1, 2, 3. Les auteurs utilisent le catalogue de Robbins et n'évaluent que les cratères de plus de 6 km. Pour la région 1, on obtient presque le même



FIGURE 6.3: Dalle 1 a), dalle 2 b) et dalle 3 c). Les cercles verts indiquent tous les véritables cratères (positifs) qui correspondent au moins à l'un des quatre catalogues. Les cercles bleus indiquent tous les cratères étiquetés manuellement comme étant des cratères. Les rectangles verts correspondent aux cratères du catalogue Barlow. Les rectangles rouges sont ceux qui sont omis par notre méthode. Les cercles rouges représentent les mauvais cratères détectés.



FIGURE 6.4: Dalle 1 a), dalle 2 b) et dalle 3 c). Les cercles verts indiquent les TP. Les cercles bleus indiquent les cratères qui ne figurent dans aucun des quatre catalogues. Les rectangles verts correspondent aux cratères du catalogue de Robbins. Les rectangles rouges sont les FN. Les cercles rouges représentent les FP.

	Notre approche			App	proche de	Di
	D (%)	В	Q (%)	D (%)	В	Q (%)
Site 1	76.3	0.08	71.7	76	0.16	68
Site 2	81.5	0.18	71.2	76	0.14	69
Site 3	96.1	0.23	78.6	90	0.15	79

TABLE 6.5: Comparaison avec l'approche de détection de
Di [41].

	Notre approche			App	roche de	Bue
	D (%)	В	Q (%)	D (%)	В	Q (%)
Site 1	85.8	0.16	75.6	76	0.29	61
Site 2	92.2	0.29	73.0	non	non	non
Site 3	98.1	0.39	71.1	non	non	non

TABLE 6.6: Comparaison avec l'approche de détection deBue et Stepinski [17].

taux de détection que Di *et al.*, mais nos détections ont un facteur de qualité plus élevé. Pour les régions 2 et 3, on a toujours de meilleurs résultats que Di *et al.* (de l'ordre de 5 - 6%).

Pour Di *et al.* [41], les valeurs TP et FP changent lorsque le seuil T de classifieur est ajusté. Di *et al.* [41] limitent le facteur d'évaluation B à environ 0.15 et ne peuvent pas extraire de nombreux cratères. La plupart de ces cratères sont profondément dégradés et les cratères modifiés sont rejetés par le classifieur. Notamment, le plus petit cratère détecté avec une fenêtre 20×20 pixels est d'environ 4×4 pixels. Les très petits cratères ne peuvent pas être détectés par leur approche à cause des caractéristiques du classifieur.

On compare aussi les résultats de Di avec les facteurs de qualité que l'on obtient avant la délimitation précise. Pour les trois régions, nous obtenons un meilleur pourcentage de détection et un facteur de qualité plus élevé pour la région 1 et la région 2. Pour la région 1, nous avons le même facteur de branchement ("branching factor") que Di *et al.*; cependant, pour les régions 2 et 3, Di *et al.* obtiennent un meilleur facteur de branchement.

La comparaison avec l'approche de Bue et Stepinski peut être effectuée seulement sur la région 1. Pour la validation, ils utilisent le catalogue de Barlow, tout comme nous le faisons. Pour la région 1, nous obtenons de meilleurs des résultats que l'approche de Bue et Stepinski [17], qui ont atteint D = 76%, B = 0.29 et Q = 61%. On doit noter que : 1) les résultats de Bue et Stepinski [17] sont obtenus en utilisant le catalogue de Barlow; 2) dans le cas de l'approche de Bue et Stepinski, il est aussi très compliqué de détecter des cratères dégradés.

6.3 Conclusion

Dans ce chapitre, nous avons détaillé et évalué notre algorithme de détection des cratères d'impact. Nous avons comparé nos résultats avec deux catalogues construit manuellement. Cette étude nous a permis d'estimer les facteurs de qualité qui nous ont servi à nous comparer avec les méthodes de Di *et al.* et Bue et Stepinski. La détection varie selon la complexité des régions testées. Les statistiques montrent que notre algorithme fournit des bons résultats par rapport aux autres approches. On peut noter que notre approche pourrait être optimisée pour obtenir une meilleure détection des gros cratères (D > 25 km).

Chapitre 7

Conclusion et perspectives

Sommaire

7.1	Synthèse	79
7.2	Contributions	80
7.3	Perspectives	80

7.1 Synthèse

En planétologie, la détermination de l'âge de la surface des différents corps célestes du système solaire est l'un des grands défis de l'astrophysique. La datation précise des surfaces est en effet fondamentale pour comprendre les processus de formation et d'évolution des différents objets. À partir d'hypothèses sur le flux d'impacteurs, la mesure de la distribution des cratères d'impact permet d'estimer l'âge de la surface impactée. En raison de l'énorme quantité de données à traiter et afin de faciliter ce processus de datation, des approches automatiques ont été proposées pour détecter automatiquement les cratères d'impact. Dans les algorithmes proposés, les valeurs de couleur des images ou les valeurs d'altitude du modèle numérique d'élévation sont généralement utilisées.

Dans cette thèse, nous avons proposé une méthode basée sur l'apprentissage automatique, tout comme Di *et al.* [41] en 2014, en combinant expertise et automatisation. Comme Schmidt *et al.* [129], notre méthode travaille en 3D et donne des informations sur la distribution des cratères, comme dans Chapman [22]. De plus, l'approche ne repose pas uniquement sur les informations concernant la topographie, mais également sur la courbure en chacun des sommets, c'est-à-dire que l'on prend en considération la concavité forte/faible ou la convexité en chaque sommet, comme dans Schmidt *et al.* [129]. Les courbures sont connues en modélisation géométrique pour donner des informations importantes sur la géométrie [77].

L'approche présentée ici tente donc d'utiliser une méthode combinée manuelle/experte, qui fonctionne en 3D, et qui utilise toutes les informations possibles sur la topographie. Elle produit non seulement les fréquences mais aussi des diamètres, les morphologies et des positions. Nous avons comparé les performances de nos approches avec celles de Di *et al.*[41] et de Bue et Stepinski [17], car il s'agit des approches qui se rapprochent le plus des travaux proposés dans cette thèse. Une approche entièrement automatique est basée sur une pré-détection rapide des cratères qui servirait de base à la validation humaine. Notre approche tient compte de cette limitation et ne vise pas à remplacer entièrement les experts humains. Nous proposons un algorithme de détection de cratères utilisant des informations sélectionnées manuellement pour la validation afin de rendre les résultats plus fiables.

Dans cette thèse, une méthode innovante automatique est proposée pour détecter efficacement les cratères à partir de maillages 3D, mais aussi pour caractériser la morphologie des cratères détectés. Cette méthode peut également être appliquée à des DEM 2.5D classiques. La fréquence spatiale des cratères et les morphologies faciliteront à terme la datation d'une surface planétaire et la compréhension du processus de dégradation active dans le temps.

7.2 Contributions

On peut résumer les résultats de cette thèse par ces points suivants :

- A partir d'un raisonnement sur des distributions gaussiennes des valeurs des courbures, nous avons proposé un filtre simple pour diminuer l'influence des bruits "Along-Track" et "Cross-Track" (chapitre 3);
- Nous avons développé un algorithme générique de représentation des cratères en se basant sur la morphologie des cratères d'impact. Plusieurs valeurs de *P* ont été testées. Le choix parmi ces valeurs dépend du nombre de caractéristiques (niveaux) que l'on veut exprimer (chapitre 3);
- Nous avons comparé expérimentalement l'avantage de l'utilisation des courbures 3D au lieu des DEM avec quatre caractéristiques pour l'extraction des cratères : pseudo-Haar, pseudo-Haar mis en échelle, LBP, en utilisant de six classifieurs (chapitre 3 et chapitre 4);
- Nous avons comparé expérimentalement l'aptitude de description des caractéristiques des cratères : pseudo-Haar, pseudo-Haar mis en échelle, LBP, à l'aide de six classifieurs (chapitre 4);
- Nous avons comparé expérimentalement la performance d'apprentissage de six classifieurs : k plus proches voisins, séparateurs à vastes marges, arbres de décision, réseau de neurones, "bagging tree" et "boosting" sur huit bases de données (chapitre 4);
- Nous avons fait une analyse expérimentale des cratères, en les catégorisant dans trois classes différentes : valide, secondaire, dégradé (chapitre 5);
- Nous avons comparé les résultats de notre algorithme de détection avec deux catalogues [9, 120], afin de pouvoir se comparer avec les approches de Di *et al.* [41] et de Bue et Stepinski [17] (chapitre 6).

7.3 Perspectives

L'algorithme de description des cratères en multi-niveaux offre beaucoup de possibilités d'optimisation et de travail futur. Le paramètre *P* peut toujours être amélioré avec des propositions de nouvelles stratégies d'optimisation. Une voie de développement futur sera d'étudier en détails les morphologies d'autres surfaces planétaires.

Notre approche a montré une bonne détection des cratères de taille inférieure à 6 km. Cependant, il n'est encore pas possible d'extraire des cratères plus petits que 1.5 - 2 km. Cela nous encourage à lancer notre algorithme sur des données de plus grandes résolutions.

Les résultats prometteurs obtenus en détection des bords des cratères nous encouragent à poursuivre l'extraction d'autres caractéristiques des cratères.

Les bons résultats obtenus en catégorisation des cratères nous encouragent à poursuivre l'étude sur de plus grandes bases de cratère, de faire des nouvelles catégories et d'essayer de se positionner parmi les autres études faites sur ce sujet.

Bibliographie

- O. Abramov and S. J. Mojzsis. Thermal effects of impact bombardments on Noachian Mars. *Earth and Planetary Science Letters*, 442 :108– 120, 2016.
- [2] Khan Academy. Contour maps, 2018.
- [3] D. L. Anderson, W. F. Miller, G. V. Latham, Y. Nakamura, M. N. Toksöz, A. M. Dainty, F. K. Duennebier, A. R. Lazarewicz, R. L. Kovach, and T. C. D. Knight. Seismology on Mars. *Journal of Geophysical Research*, 82(28) :4524–4546, 1977.
- [4] R. Arvidson, E. Guinness, and S. Lee. Differential aeolian redistribution rates on Mars. *Nature*, 278:533–535, 1979.
- [5] L. Bandeira, W. Ding, and T. F. Stepinski. Detection of sub-kilometer craters in high resolution planetary images using shape and texture features. *Advances in Space Research*, 49(1):64–74, 2012.
- [6] L. Bandeira, J. Saraiva, and P. Pina. Impact Crater Recognition on Mars Based on a Probability Volume Created by Template Matching. *IEEE Transactions on Geoscience and Remote Sensing*, 45(12):4008–4015, 2007.
- [7] T. Barata, E. I. Alves, A. Machado, and G. A. Barberes. Characterization of palimpsest craters on Mars. *Planetary and Space Science*, 72(1):62–69, 2012.
- [8] T. Barata, E. I. Alves, J. Saraiva, and P. Pina. Automatic Recognition of Impact Craters on the Surface of Mars. In Aurélio Campilho and Mohamed Kamel, editors, *Image Analysis and Recognition*, pages 489– 496, Berlin, Heidelberg, 2004. Springer Berlin Heidelberg.
- [9] N. G. Barlow. *Relative ages and the geologic evolution of Martian terrain units*. PhD thesis, Arizona Univ., Tucson., 1987.
- [10] N. G. Barlow. Crater size-frequency distributions and a revised Martian relative chronology. *Icarus*, 75(2) :285–305, 1988.
- [11] N. G. Barlow. The degradation of impact craters in Maja Valles and Arabia, Mars. *Journal of Geophysical Research : Planets*, 100(E11) :23307– 23316, 1995.
- [12] N. G. Barlow. Impact craters in the northern hemisphere of Mars : Layered ejecta and central pit characteristics. *Meteoritics & Planetary Science*, 41(10) :1425–1436, 2006.
- [13] N. G. Barlow. Revision of the "Catalog of large martian impact craters" and comparison to other martian crater databases. *Lunar and Planetary Science XLVIII*, 2017.
- [14] C. M. Bishop. *Pattern Recognition and Machine Learning*. Springer, 2006.
- [15] P. A. Bland and N. A. Artemieva. Efficient disruption of small asteroids by Earth's atmosphere. *Nature*, 424 :288—-291, 2003.

- [16] J. P. Boyd. Scientific Visualization and Information Architecture, Chapter 4 : Contour Plots, Course Notes, 2000.
- [17] B. D. Bue and T. F. Stepinski. Machine Detection of Martian Impact Craters From Digital Topography Data. *IEEE Transactions on Geoscience and Remote Sensing*, 45(1):265–274, 2007.
- [18] M. C. Burl, T. Stough, W. Colwell, E. B. Bierhaus, W. J. Merline, and C. Chapman. Automated detection of craters and other geological features. 6th International Symposium on Artifical Intelligence, Robotics and Automation in Space; Montreal, Quebec; Canada, 2001.
- [19] C. Capanna, L. Jorda, P. Lamy, and G. Gesquiere. 3D Reconstruction of small solar system bodies using photoclinometry by deformation. *IADIS International Journal on Computer Science and Information Systems*, 7(1) :32–46, 2012.
- [20] C. F. Carlson. Lecture 10 : Hough circle transform. Rochester Institute of Technology : Lecture Notes, 2005.
- [21] M. P. D. Carmo. *Differential Geometry of Curves and Surfaces*. Prentice-Hall, 1976.
- [22] C. R. Chapman. Cratering on asteroids from Galileo and NEAR Shoemaker. Asteroids III, pages 315–329, 2002.
- [23] C. R. Chapman, W. J. Merline, P. C. Thomas, J. Joseph, A. F. Cheng, and N. Izenberg. Impact History of Eros : Craters and Boulders. *Icarus*, 155(1) :104–118, 2002.
- [24] M. Chen, M. Lei, D. Liu, Y. Zhou, H. Zhao, and K. Qian. Morphological Features-Based Descriptive Index System for Lunar Impact Craters. *ISPRS International Journal of Geo-Information*, 7(1), 2018.
- [25] Y. Cheng, A. E. Johnson, L. H. Matthies, and C. F. Olson. Optical landmark detection for spacecraft navigation. In *Proceedings of the* 13th Annual AAS/AIAA Space Flight Mechanics Meeting, 2003.
- [26] N. Christoff. Feature extraction and classification using minimal curvature of 3D mesh for automatic crater detection. In 52nd International Scientific Conference on Information, Communication and Energy Systems and Technologies (ICEST2017), pages 44–47, 2017.
- [27] N. Christoff, L. Jorda, S. Viseur, S. Bouley, A. Manolova, and J.-L. Mari. Discrete curvatures combined with machine learning for automated extraction of impact craters on 3D topographic meshes. In EGU General Assembly Conference Abstracts, volume 19 of EGU General Assembly Conference Abstracts, page 8549, 2017.
- [28] N. Christoff, A. Manolova, L. Jorda, and J.-L. Mari. Feature extraction and automatic detection of martian impact craters from 3D meshes. In 2017 13th International Conference on Advanced Technologies, Systems and Services in Telecommunications (TELSIKS), pages 211–214, 2017.
- [29] N. Christoff, A. Manolova, L. Jorda, and J.-L. Mari. Morphological Crater Classification via Convolutional Neural Network with Application on MOLA data. In ANNA '18; Advances in Neural Networks and Applications 2018, pages 1–5, 2018.
- [30] N. Christoff, A. Manolova, L. Jorda, S. Viseur, S. Bouley, and J.-L. Mari. Level-Set Based Algorithm for Automatic Feature Extraction

on 3D Meshes : Application to Crater Detection on Mars. In *Computer Vision and Graphics*, pages 103–114, Cham, 2018. Springer International Publishing.

- [31] W. S. Cleveland, S. J. Devlin, and E. Grosse. Regression by local fitting : Methods, properties, and computational algorithms. *Journal of Econometrics*, 37(1):87–114, 1988.
- [32] D.-A. Clevert, T. Unterthiner, and S. Hochreiter. Fast and Accurate Deep Network Learning by Exponential Linear Units (ELUs). *CoRR*, abs/1511.07289, 2015.
- [33] J. P. Cohen and W. Ding. Crater detection via genetic search methods to reduce image features. *Advances in Space Research*, 53(12) :1768– 1782, 2014.
- [34] J. P. Cohen, H. Z. Lo, T. Lu, and W. Ding. Crater Detection via Convolutional Neural Networks. *CoRR*, abs/1601.00978, 2016.
- [35] F. M. Costard. The spatial distribution of volatiles in the Martian hydrolithosphere. *Earth, Moon, and Planets*, 45(3) :265–290, 1989.
- [36] G. C. H. E. De Croon, D. Izzo, and G. Schiavone. Time-to-Contact Estimation in Landing Scenarios using Feature Scales. *Acta Futura*, 5:73–82, 2012.
- [37] J. C. Davis. Statistics and Data Analysis in Geology. John Wiley & Sons, Inc., New York, NY, USA, 2nd edition, 1990.
- [38] B. de Cambray and T s. Yeh. A Multidimensional (2D, 2.5D, and 3D) Geographical Data Model. In *in Proceedings of the Sixth International Conference on Management of Data, COMAD*'94, 1994.
- [39] M. R. Dence. Dimensional analysis of impact structures. *Meteoritics*, 8:343–344, 1973.
- [40] M. R. Dence, M. J. S. Innes, and P. B. Robertson. Recent geological and geophysical studies of Canadian craters. *Shock Metamorphism of Natural Materials (B. M. French and N. M. Short, eds.)*, pages 339–362, 1968.
- [41] K. Di, W. Li, Z. Yue, Y. Sun, and Y. Liu. A machine learning approach to crater detection from topographic data. *Advances in Space Research*, 54(11):2419–2429, 2014.
- [42] M. Ding, Y. Cao, and Q. Wu. Novel approach of crater detection by crater candidate region selection and matrix-pattern-oriented least squares support vector machine. *Chinese Journal of Aeronautics*, 26(2):385–393, 2013.
- [43] W. Ding, T. F. Stepinski, Y. Mu, L. Bandeira, , R. Ricardo, Y. Wu, Z. Lu, T. Cao, and X. Wu. Subkilometer Crater Discovery with Boosting and Transfer Learning. ACM Transactions on Intelligent Systems and Technology (TIST), 2(4):1–22, 2011.
- [44] R. O. Duda, P. E. Hart, and D. G. Stork. *Pattern Classification (2Nd Edition)*. Wiley-Interscience, New York, NY, USA, 2000.
- [45] L. Ferrière. Meteor Impact on Earth : Excavation and Modification, 2014.

- [46] N. K. Forsberg-Taylor, A. D. Howard, and R. A. Craddock. Crater degradation in the Martian highlands : Morphometric analysis of the Sinus Sabaeus region and simulation modeling suggest fluvial processes. *Journal of Geophysical Research (Planets)*, 109 :E05002, 2004.
- [47] B. M. French. *Traces of catastrophe : A handbook of shock-metamorphic effects in terrestrial meteorite impact structures.* LPI Contribution No. 954, Lunar and Planetary Institute, Houston, 1998.
- [48] Y. Freund and R. E. Schapire. A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting. *Journal of Computer* and System Sciences, 55(1):119–139, 1997.
- [49] S. Gauthier, W. Puech, R. Bénière, and G. Subsol. Analysis of digitized 3D mesh curvature histograms for reverse engineering. *Computers in Industry*, 92-93 :67–83, 2017.
- [50] R. A. F Grieve and M. J. Cintala. A method for estimating the initial impact conditions of terrestrial cratering events, exemplified by its application to Brent crater, Ontario. *Proceedings of Lunar and Planetary Science Conference*, 12(B) :1607–1621, 1981.
- [51] R. A. F. Grieve, M. R. Dence, and P. B. Robertson. Cratering processes - As interpreted from the occurrence of impact melts. *Impact and Explosion Cratering : Planetary and Terrestrial Implications*, pages 791–814, 1977.
- [52] R. A. F Grieve, P. B. Robertson, and M. R. Dence. Constraints on the formation of ring impact structures, based on terrestrial data. *Multiring basins : Formation and evolution; Proceedings of the Lunar and Planetary Science Conference. New York and Oxford, Pergamon Press,* pages 37–57, 1981.
- [53] K. Gwinner, R. Jaumann, E. Hauber, H. Hoffmann, C. Heipke, J. Oberst, G. Neukum, V. Ansan, J. Bostelmann, A. Dumke, S. Elgner, G. Erkeling, F. Fueten, H. Hiesinger, N. M. Hoekzema, E. Kersten, D. Loizeau, K.-D. Matz, P. C. McGuire, V. Mertens, G. Michael, A. Pasewaldt, P. Pinet, F. Preusker, D. Reiss, T. Roatsch, R. Schmidt, F. Scholten, M. Spiegel, R. Stesky, D. Tirsch, S. van Gasselt, S. Walter, M. Wählisch, and K. Willner. The High Resolution Stereo Camera (HRSC) of Mars Express and its approach to science analysis and mapping for Mars and its satellites. *Planetary and Space Science*, 126 :93–138, 2016.
- [54] N. Haala, R. Reulke, M. Thies, and T. Aschoff. Combination of terrestrial laser scanning with high resolution panoramic images for investigations in forest applications and tree species recognition. *Conference IAPRS and SIS*, XXXIV, part 5/W16, 2004.
- [55] H. Hargitai. *Secondary Crater*, pages 1–7. Encyclopedia of Planetary Landforms, Springer New York, New York, NY, 2014.
- [56] W. K. Hartmann. Secular changes in meteoritic flux through the history of the solar system. *Icarus*, 4(2) :207–213, 1965.
- [57] W. K. Hartmann. Lunar cratering chronology. *Icarus*, 13(2):299–301, 1970.
- [58] W. K. Hartmann. Martian cratering III : Theory of crater obliteration. *Icarus*, 15(3) :410–428, 1971.
- [59] W. K. Hartmann. Paleocratering of the Moon : Review of post-Apollo data. Astrophysics and Space Science, 17(1) :48–64, 1972.
- [60] W. K. Hartmann and G. Neukum. Cratering Chronology and the Evolution of Mars. *Space Science Reviews*, 96(1):165–194, 2001.
- [61] D. M. Hawkins. The Problem of Overfitting. Journal of Chemical Information and Computer Sciences, 44(1):1–12, 2004.
- [62] K. He, X. Zhang, S. Ren, and J. Sun. Deep Residual Learning for Image Recognition. *CoRR*, abs/1512.03385, 2015.
- [63] R. Honda, Y. Iijima, and O. Konishi. Mining of Topographic Feature from Heterogeneous Imagery and Its Application to Lunar Craters. *Progress in Discovery Science : Final Report of the Japanese Dicsovery Science Project*, pages 395–407, 2002.
- [64] C.-W. Hsu, Ch.-Ch. Chang, and Ch.-J. Lin. A practical guide to support vector classification , 2010.
- [65] S. Ioffe and C. Szegedy. Batch Normalization : Accelerating Deep Network Training by Reducing Internal Covariate Shift. *CoRR*, abs/1502.03167, 2015.
- [66] G. James, D. Witten, T. Hastie, and R. Tibshirani. An Introduction to Statistical Learning with Applications in R. Springer-Verlag New York, 2013.
- [67] W. Jarmołowski and J. Łukasiak. A Study on Along-Track and Cross-Track Noise of Altimetry Data by Maximum Likelihood : Mars Orbiter Laser Altimetry (Mola) Example. *Artificial Satellites*, 50 :143–155, 2015.
- [68] S. Jin and T. Zhang. Automatic detection of impact craters on Mars using a modified adaboosting method. *Planetary and Space Science*, 99:112–117, 2014.
- [69] B. C. Johnson and H. J. Melosh. Formation of spherules in impact produced vapor plumes. *Icarus*, 217(1):416–430, 2012.
- [70] C. Kanan and G. W. Cottrell. Color-to-Grayscale : Does the Method Matter in Image Recognition? *PLOS ONE*, 7(1) :1–7, 2012.
- [71] N. Kasabov. Foundations of Neural Networks, Fuzzy Systems, and Knowledge Engineering. The MIT Press, 1996.
- [72] T. Kenkmann, G. S. Collins, and K. Wünnemann. *The Modification Stage of Crater Formation*. 2012.
- [73] J. R. Kim, J.-P. Muller, S. van Gasselt, J. G. Morley, and G. Neukum. Automated crater detection, a new tool for Mars cartography and chronology. *Photogrammetric Engineering and Remote Sensing*, 71(10):1205–1217, 2005.
- [74] T. Kneissl, S. van Gasselt, and G. Neukum. Map-projectionindependent crater size-frequency determination in GIS environments—New software tool for ArcGIS. *Planetary and Space Science*, 59(11):1243–1254, 2011.
- [75] S. O. Krøgli and H. Dypvik. Automatic detection of circular outlines in regional gravity and aeromagnetic data in the search for impact structure candidates. *Computers & Geosciences*, 36(4):477–488, 2010.

- [76] D. Kudelski. Détection automatique d'objets géologiques à partir de données numériques d' affleurements 3D. PhD thesis, Université de Provence, 2011.
- [77] D. Kudelski, S. Viseur, G. Scrofani, and J.-L. Mari. Feature Line Extraction on Meshes through Vertex Marking and 2D Topological Operators. *International Journal of Image and Graphics*, 11(04):531–548, 2011.
- [78] M. Küppers, R. Moissl, J.-B. Vincent, S. Besse, S. F. Hviid, B. Carry, B. Grieger, H. Sierks, H. U. Keller, S. Marchi, OSIRIS Team, M. A'Hearn, F. Angrilli, C. Barbieri, A. Barucci, J.-L. Bertaux, G. Cremonese, V. Da Deppo, B. Davidsson, S. Debei, M. De Cecco, S. Fornasier, M. Fulle, O. Groussin, P. Gutiérrez, L. Koschny D. W.-H. Ip, Jorda, J. Knollenberg, J. R. Kramm, E. Kührt, P. Lamy, L. M. Lara, M. Lazzarin, J. J. López-Moreno, F. Marzari, H. N. Michalik, G., H. Rickman, R. Rodrigo, L. Sabau, N. Thomas, and K.-P. Wenzel. Boulders on Lutetia. *Planetary and Space Science*, 66 :71–78, 2012.
- [79] R. O. Kuzmin, N. N. Bobina, E. V. Zabalueva, and V. P. Shashkina. Structure inhomogeneities of the Martian cryosphere. *Solar System Research*, 22 :121–133, 1988.
- [80] A. Lagain. Datation et caractérisation morphologique des cratères à éjecta lobés des terrains de l'hémisphère Nord martien et variation du flux d'impact météoritique. Master's thesis, Université Paris-Sud (Orsay), 2014.
- [81] A. Lagain and C. Marmo. alagain/martian_crater_database 1.1, 2018.
- [82] B. Leroy, G. Medioni, E. Johnson, and L. Matthies. Crater detection for autonomous landing on asteroids. *Image and Vision Computing*, 19(11):787–792, 2001.
- [83] B. Li, Z. Ling, J. Zhang, and Z. Wu. Automatic Detection and Boundary Extraction of Lunar Craters Based on LOLA DEM Data. *Earth*, *Moon, and Planets*, 115(1):59–69, 2015.
- [84] S. Li, P. Cui, and H. Cui. Autonomous navigation and guidance for landing on asteroids. *Aerospace Science and Technology*, 10(3):239–247, 2006.
- [85] M. M. Lipschultz. Theory and Problems of Differential Geometry : Schaum's Outline. McGraw-Hill, New York, 1969.
- [86] S. Liu, W. Ding, J. P. Cohen, D. A. Simovici, and T. F. Stepinski. Bernoulli trials based feature selection for crater detection. In GIS, 2011.
- [87] L. Luo, X. Y. Wang, W. Ji, and C. Li. Automated detection of lunar craters based on Chang'E-1 CCD data. *Image and Signal Processing* (CISP), 2:883–887, 2011.
- [88] B. Maass, H. Krüger, and S. Theil. An edge-free, scale-, pose- and illumination-invariant approach to crater detection for spacecraft navigation. In 2011 7th International Symposium on Image and Signal Processing and Analysis (ISPA), pages 603–608, 2011.
- [89] M. C. Malin, J. F. Bell, B. A. Cantor, M. A. Caplinger, W. M. Calvin, R. T. Clancy, K. S. Edgett, L. Edwards, R. M. Haberle, P. B. James, S. W. Lee, M. A. Ravine, P. C. Thomas, and M. J. Wolff. Context Camera Investigation on board the Mars Reconnaissance Orbiter. *Journal of Geophysical Research : Solid Earth*, 112(5), 2007.

- [90] M. C. Malin, K. S. Edgett, B. A. Cantor, M. A. Caplinger, G. E. Danielson, E. H. Jensen, M. A. Ravine, J. L. Sandoval, and K. D. Supulver. An overview of the 1985-2006 Mars Orbiter Camera science investigation. *International Journal of Mars Science and Exploration*, 5 :1–60, 2010.
- [91] A. Manolova. *Catégorisation par mesures de dissimilitude et caractérisation d'images en multi échelle*. PhD thesis, Thèse en co tutelle : L'université de Grenoble et L'université Technique de Sofia, Bulgarie, 2011.
- [92] S. Marchi, M. Massironi, J.-B. Vincent, A. Morbidelli, S. Mottola, F. Marzari, M. Küppers, S. Besse, N. Thomas, C. Barbieri, G. Naletto, and H. Sierks. The cratering history of asteroid (21) Lutetia. *Planetary and Space Science*, 66(1):87–95, 2012.
- [93] R. Martins, P. Pina, J. S. Marques, and M. Silveira. Crater detection by a boosting approach. *IEEE Geoscience and Remote Sensing Letters*, 6(1):127–131, 2009.
- [94] D. E. Maxwell. Simple Z model for cratering, ejection, and the overturned flap. *Impact and Explosion Cratering : Planetary and Terrestrial Implications*, pages 1003–1008, 1977.
- [95] A. S. McEwen and E. B. Bierhaus. The importance of secondary cratering to age constraints on planetary surfaces. *Annual Review of Earth and Planetary Sciences*, 34(1):535–567, 2006.
- [96] H. J. Melosh. Impact cratering : A geologic process. 1989.
- [97] H. J. Melosh and B. Ivanov. Impact Crater Collapse. *Annual Review of Earth and Planetary Sciences*, 27(1):385–415, 1999.
- [98] T. Michikami, A. M. Nakamura, and N. Hirata. The shape distribution of boulders on Asteroid 25143 Itokawa : Comparison with fragments from impact experiments. *Icarus*, 207(1) :277–284, 2010.
- [99] M. F. Møller. A scaled conjugate gradient algorithm for fast supervised learning. *Neural Networks*, 6(4) :525–533, 1993.
- [100] I. Mynatt, S. Bergbauer, and D. D. Pollard. Using differential geometry to describe 3-D folds. *Journal of Structural Geology*, 29(7) :1256– 1266, 2007.
- [101] G. Neukum, B. A. Ivanov, and W. K. Hartmann. Cratering Records in the Inner Solar System in Relation to the Lunar Reference System. *Space Science Reviews*, 96(1):55–86, 2001.
- [102] G. Neukum and R. Jaumann. HRSC : the High Resolution Stereo Camera of Mars Express. In A. Wilson and A. Chicarro, editors, *Mars Express : the Scientific Payload*, volume 1240 of *ESA Special Publication*, pages 17–35, 2004.
- [103] G. Neukum, B. König, and J. Arkani-Hamed. A study of lunar impact crater size-distributions. *The moon*, 12(2) :201–229, 1975.
- [104] G. Neukum, B. König, H. Fechtig, and D. Storzer. Cratering in the earth-moon system-Consequences for age determination by crater counting. *Proceedings Lunar Science Conference*, 6th, Houston, 3 :2597– 2620, 1975.
- [105] T. Ojala, M. Pietikäinen, and D. Harwood. A comparative study of texture measures with classification based on featured distributions. *Pattern Recognition*, 29(1):51–59, 1996.

- [106] J. D. O'Keefe and T. J. Ahrens. Cometary and meteorite swarm impact on planetary surfaces. *Journal of Geophysical Research : Solid Earth*, 87(B8) :6668–6680, 1982.
- [107] J. D. O'Keefe and T. J. Ahrens. Planetary cratering mechanics. *Journal of Geophysical Research : Planets*, 98(E9) :17011–17028, 1993.
- [108] J. Ormö and M. Lindström. When a cosmic impact strikes the sea bed. *Geological Magazine*, 137(1):67–80, 2000.
- [109] S. Osher and J. A. Sethian. Fronts propagating with curvaturedependent speed : Algorithms based on Hamilton-Jacobi formulations. *Journal of Computational Physics*, 79(1):12–49, 1988.
- [110] L. F. Palafox, C. W. Hamilton, S. P. Scheidt, and A. M. Alvarez. Automated detection of geological landforms on Mars using Convolutional Neural Networks. *Computers and Geosciences*, 101:48–56, 2017.
- [111] Q. R. Passey and H. J. Melosh. Effects of atmospheric breakup on crater field formation. *Icarus*, 42(2) :211–233, 1980.
- [112] M. M. Pedrosa, S. C. de Azevedo, E. A. da Silva, and M. A. Dias. Improved automatic impact crater detection on Mars based on morphological image processing and template matching. *Geomatics, Natural Hazards and Risk*, 8(2) :1306–1319, 2017.
- [113] R. J. Pike. Control of crater morphology by gravity and target type -Mars, earth, moon. *Lunar and Planetary Science Conference Proceedings*, 11:2159–2189, 1980.
- [114] A. Polette, J. Meunier, and J.-L. Mari. "Shape-Curvature-Graph" : Towards a New Model of Representation for the Description of 3D Meshes. In Augmented Reality, Virtual Reality, and Computer Graphics -4th International Conference, AVR 2017, Ugento, Italy, June 12-15, 2017, Proceedings, Part II, In Lecture Notes in Computer Science, LNCS 10325, Springer, pages 369–384, 2017.
- [115] M. Ponti, T. S. Nazaré, and G. S. Thumé. Image quantization as a dimensionality reduction procedure in color and texture feature extraction. *Neurocomputing*, 173:385–396, 2016.
- [116] I. M. Pritchard. Automated Impact Crater Detection and Characterization Using Digital Elevation Data. PhD thesis, The University of Western Ontario, 2016.
- [117] R. J. Quinlan. C4.5: Programs for Machine Learning. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1993.
- [118] D. G. Rees. Foundations of Statistics. CRC Press, pages 244–249, 1987.
- [119] B. D. Ripley. Cambridge : Cambridge University, 1996.
- [120] S. J. Robbins and B. M. Hynek. A new global database of Mars impact craters ≥ 1 km : 1. Database creation, properties, and parameters. *Journal of Geophysical Research : Planets*, 117(E5), 2012.
- [121] S. J. Robbins and B. M. Hynek. The secondary crater population of Mars. *Earth and Planetary Science Letters*, 400 :66–76, 2014.
- [122] D. J. Roddy, N. R. Isbell, C. L. Mardock, T. M. Hare, M. B. Wyatt, L. M. Soderblom, and J. M. Boyce. I. Martian impact craters, ejecta blankets, and related morphologic features : Computer digital inventory in Arc/Info and Arcview format. *Lunar and Planetary Science XXIX*, page abstract 1874, 1998.

- [123] J. F. Rodionova, K. I. Dekchtyareva, A. A. Khramchikhin, G. G. Michael, S. V. Ajukov, S. G. Pugacheva, and V. V. Shevchenko. *Morpholo*gical Catalogue Of The Craters Of Mars. ESA/ESTEC Noordwijk, 2000.
- [124] G. Salamunićcar and S. Lončarić. Open framework for objective evaluation of crater detection algorithms with first test-field subsystem based on MOLA data. *Advances in Space Research*, 42(1):6–19, 2008.
- [125] G. Salamunićcar and S. Lončarić. Application of machine learning using support vector machines for crater detection from Martian digital topography data. 38th COSPAR Scientific Assembly. Held 18-15 July 2010, in Bremen, Germany, page 3, 2010.
- [126] G. Salamunićcar, S. Lončarić, A. Grumpe, and C. Wóhler. Hybrid method for crater detection based on topography reconstruction from optical images and the new LU78287GT catalogue of Lunar impact craters. *Advances in Space Research*, 53(12) :1783 – 1797, 2014.
- [127] G. Salamunićcar, S. Lončarić, and E. Mazarico. LU60645GT and MA132843GT catalogues of Lunar and Martian impact craters developed using a Crater Shape-based interpolation crater detection algorithm for topography data. *Planetary and Space Science*, 60(1):236–247, 2012.
- [128] Y. Sawabe, T. Matsunaga, and S. Rokugawa. Automated detection and classification of lunar craters using multiple approaches. *Advances in Space Research*, 37(1):21–27, 2006.
- [129] M. P. Schmidt, J. Muscato, S. Viseur, L. Jorda, S. Bouley, and J.-L. Mari. Robust Detection of Round Shaped Pits Lying on 3D Meshes Application to Impact Crater Recognition. *EGU General Assembly* 2015, 17(EGU2015) :7628, 2015.
- [130] E. M. Shoemaker. CHAPTER 8 Interpretation of Lunar Craters †. In Zdeněk Kopal, editor, *Physics and Astronomy of the Moon*, pages 283– 359. Academic Press, 1961.
- [131] J. A. Shufelt. Performance evaluation and analysis of monocular building extraction from aerial imagery. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21(4):311–326, 1999.
- [132] D. E. Smith, M. T. Zuber, H. V. Frey, J. B. Garvin, J. W. Head, D. O. Muhleman, G. H. Pettengill, R. J. Phillips, S. C. Solomon, H. J. Zwally, W. B. Banerdt, and T. C. Duxbury. Topography of the Northern Hemisphere of Mars from the Mars Orbiter Laser Altimeter. *Science*, 279(5357):1686–1692, 1998.
- [133] T. F. Stepinski, W. Ding, and R. Vilalta. Detecting Impact Craters in Planetary Images Using Machine Learning. *Intelligent Data Analysis* for Real-Life Applications : Theory and Practice, pages 146–159, 2012.
- [134] S. A. Stewart. Estimates of yet-to-find impact crater population on Earth. *Journal of the Geological Society*, 168(1) :1–14, 2011.
- [135] D. Stöffler, D. E. Gault, J. Wedekind, and G. Polkowski. Experimental hypervelocity impact into quartz sand : Distribution and shock metamorphism of ejecta. *Journal of Geophysical Research*, 80(29) :4062–4077, 1975.

- [136] B. Stout. Chapter 1 Systèmes de coordonnées. http ://www.fresnel.fr/perso/stout/electromag/Coordonnees_ curviligne.pdf.
- [137] T. Szabó, G. Domokos, J. P. Grotzinger, and D. J. Jerolmack. Reconstructing the transport history of pebbles on Mars. *Nature communications*, 6 :8366, 2015.
- [138] K. L. Tanaka and G. J. Leonard. Geology and landscape evolution of the Hellas region of Mars. *Journal of Geophysical Research : Planets*, 100(E3) :5407–5432, 1995.
- [139] B. Trey. Enregistrement du flux d'impact par les cratères rayés lunaires : conséquences pour les datations des surfaces planétaires. PhD thesis, Ecole Doctoral Physique et Astrophysique de Lyon, 2013.
- [140] G. Troglio, J. Le Moigne, J. A. Benediktsson, G. Moser, and S. B. Serpico. Automatic Extraction of Ellipsoidal Features for Planetary Image Registration. *IEEE Geoscience and Remote Sensing Letters*, 9(1):95–99, 2012.
- [141] E. P. Turtle, E. Pierazzo, G. S. Collins, G. R. Osinski, H. J. Melosh, J. V. Morgan, W. U. Reimold, and J. G. Spray. Impact Structures : What Does Crater Diameter Mean? In *Lunar and Planetary Science Conference*, volume 35 of *Lunar and Planetary Inst. Technical Report*, 2004.
- [142] E. R. Urbach and T. F. Stepinski. Automatic detection of sub-km craters in high resolution planetary images. *Planetary and Space Science*, 57(7):880–887, 2009.
- [143] J. van Zyl. Introduction to the Physics of Remote Sensing : Imaging geometries.
- [144] S. Vijayan, K. Vani, and S. Sanjeevi. Crater detection, classification and contextual information extraction in lunar images using a novel algorithm. *Icarus*, 226(1):798–815, 2013.
- [145] T. Vinogradova, M. Burl, and E. Mjolsness. Automatic detection of sub-km craters in high resolution planetary images. *Proceedings*, *IEEE Aerospace Conference*, 7 :3201–3211, 2002.
- [146] P. Viola and M. Jones. Rapid object detection using a boosted cascade of simple features. In *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001*, volume 1, pages I–I, 2001.
- [147] J. Wang, W. Ding, B. Fradkin, C. H. Pham, P. Sherman, B. D. Tran, D. Wang, Y. Yang, and T. F. Stepinski. Effective classification for crater detection : A case study on Mars. 9th IEEE International Conference on Cognitive Informatics (ICCI'10), 2010.
- [148] Y. Wang, G. Yang, and L. Guo. A novel sparse boosting method for crater detection in the high resolution planetary image. *Advances in Space Research*, 56(5) :982–991, 2015.
- [149] S. C. Werner and T. H. Torsvik. Downsizing the Mjølnir impact structure, Barents Sea, Norway. *Tectonophysics*, 483(3):191–202, 2010.
- [150] P. G. Wetzler, R. Honda, B. Enke, W. J. Merline, C. R. Chapman, and M. C. Burl. Learning to Detect Small Impact Craters. 2005 Seventh IEEE Workshops on Applications of Computer Vision (WACV/MOTION'05), 1:178–184, 2005.

- [151] K. Wünnemann, G. S. Collins, and R. Weiss. Impact of a cosmic body into Earth's ocean and the generation of large tsunami waves : Insight from numerical modeling. *Reviews of Geophysics*, 48(4) :RG4006, 2010.
- [152] S. Zagoruyko and N. Komodakis. Wide Residual Networks. CoRR, abs/1605.07146, 2016.
- [153] T. Zhang and S. Jin. Automatic Recognition of Impact Craters on the Martian Surface from DEM and Images, pages 101–118. Springer Berlin Heidelberg, Berlin, Heidelberg, 2015.
- [154] M. T. Zuber, D. E. Smith, S. C. Solomon, D. O. Muhleman, J. W. Head, J. B. Garvin, J. B. Abshire, and J. L. Bufton. The Mars Observer laser altimeter investigation. *Journal of Geophysical Research : Planets*, 97(E5) :7781–7797, 1992.