

N° d'ordre :

N° de série :

République Algérienne Démocratique Et Populaire

Ministère De L'Enseignement Supérieur Et De La Recherche Scientifique



UNIVERSITE ECHAHID HAMMA LAKHDAR D'EL-OUED

FACULTÉ DES SCIENCES EXACTES

DEPARTEMENT D'INFORMATIQUE



MEMOIRE DE FIN D'ETUDE
MASTER ACADEMIQUE

Domaine: Mathématiques et Informatique

Filière: Informatique

Spécialité: Systèmes Distribués et Intelligence Artificiel (SDIA)

THEME

Un système d'indexation et de recherche automatique
d'images par le contenu : Cas de séisme.

Présenté par:

M^{elle}. LABBI Nour El_Hoda

M^{elle}. DAHEM Nabila

Soutenu le 03 Juin 2018

Devant le jury composé de :

Mr. GHERBI Kaddour	MA(A) l'Université d'El Oued	Encadreur
Mr. BERDJOUH Chafik	MA(B) l'Université d'El Oued	Président
Mr. LEJDEL Ibrahim	MA(A) l'Université d'El Oued	Examineur

Année universitaire: 2017/2018

Remerciements

Nous remercions Allah le tout puissant, qui nous a donné la force et la patience pour l'accomplissement de ce travail.

Nos remerciements, les plus vifs, notre profonde gratitude et nos respects s'adressent à notre Encadreur

M. GHERBI Kaddour

Pour avoir accepté de nous encadrer, pour les conseils et orientations tant précieux qu'il nous avons prodigués durant ce Mémoire.

Sans son aide, notre travail n'aurait pas vu la lumière.

Nous remercions le Dr. Settou Trablesse pour son aide, son soutien et ses conseils et nous lui souhaitons beaucoup de succès et de succès.

Nous remercions vivement les membres du jury qui nous ont fait l'honneur

D'accepter de juger notre travail.

Notre reconnaissance va aussi à tous ceux qui ont collaboré à notre Formation en particulier les enseignants du département

D'Informatique, de l'université Hama Lakhdar d'El-Oued.

Aussi à nos familles DAHEM et LABBI

Nous remercions également tous ceux qui ont participé de près Ou de loin à élaborer ce travail.

Résumé:

La classification des bâtiments endommagés est actuellement un sujet pertinent dans l'évaluation et la gestion des catastrophes, et la détection des bâtiments endommagés.

Dans ce mémoire, l'objectif est d'introduire une technique de mots visuels efficaces basée sur des descripteurs des caractéristiques (SURF) et d'Histogramme de Gradient Orienté (HOG). HOG est utilisé pour extraire des caractéristiques globales, tandis que SURF est utilisé pour l'extraction de caractéristiques locales. Nous préférons utiliser des fonctionnalités locales pour la classification et la récupération d'images à grande échelle, car il fournit une meilleure précision. La technique proposée basée sur des mots visuels donne une précision de classification de plus de 60% en utilisant la machine à vecteurs de support SVM.

Mots clés : Classification, caractéristiques locales, caractéristiques globales, SURF, HOG, BoVW, SVM.

Abstract:

The classification of damaged buildings is currently a relevant topic in disaster assessment and management, and the detection of damaged buildings .

In this memory, the objective is to introduce a technique effective visual words based on descriptors of the features (SURF) and Histogram of Oriented Gradient (HOG). HOG is used to extract global characteristics, while SURF is used for extracting local features. It prefers to use local features for large-scale image classification and retrieval because it provides better precision. The proposed technique based on visual words gives a classification accuracy of more than 60% using the SVM support vector machine.

Keywords: Classification, local characteristics , global characteristics, SURF, HOG, BoVW, SVM

المخلص:

يعتبر تصنيف المباني التالفة حالياً موضوعاً ذا صلة في تقييم وإدارة الكوارث، والكشف عن المباني المتضررة. في هذه المذكرة، الهدف هو تقديم تقنية الكلمات المرئية القائمة على استخراج الميزة (SURF) و التدرج البياني الموجه (HOG) يستخدم لاستخراج الخصائص العالمية، في حين يستخدم SURF لاستخراج الميزات المحلية. يفضل استخدام الميزات المحلية لتصنيف واسترجاع الصور واسعة النطاق لأنها توفر أفضل دقة. التقنية المقترحة على أساس الكلمات المرئية تعطي دقة التصنيف أكثر من 60% باستخدام آلة المتجهات SVM .

الكلمات المفتاحية : التصنيف ، الخصائص المحلية ، الخصائص العالمية ، SURF ، HOG ، BoVW ، SVM .

Table des matières

Table des matières	I
Liste des figures	IV
Liste des tableaux.....	VI
Liste d'abréviation	VII
Introduction générale	VIII

Chapitre I : Détection de dommage basé sur les images télédétection

Introduction	02
1. Technologies de la télédétection	02
2. Résolutions	02
2.1. La résolution spatiale	03
2.2. La résolution spectrale	03
2.3. La résolution temporelle	03
3. Acquisition des données de télédétection	04
3.1. Les Images Satellites	04
3.2. Les Photographies Aériennes	04
3.3. Comparaison entre photographies aériennes et images satellites	05
4. Domaine d'applications	05
4.1. Extraction des bâtiments.....	05
4.2. Détection de dommages.....	06
4.3. Extraction des réseaux routiers	06
4.4. La modernisation de l'agriculture	06
5. Détection et évaluation des dommages après le séisme:.....	07
5.1. Recherche sur le terrain	07
5.2. Détection de dommage par les images de télédétection	07

a.	Téledétection et dommage grades	08
b.	Dommage grades	08
5.3.	Méthode de détection damage	09
a.	Interprétation visuelle	09
b.	Techniques Multi-temporelles	10
c.	Techniques Mono-temporelles	10
	Conclusion	11

Chapitre II: Extraction des caractéristiques d’images

	Introduction.....	13
1.	Extraction des caractéristiques.....	13
1.1.	Caractéristiques globales.....	13
1.1.1.	Descripteurs de la couleur.....	13
a.	L’espace de couleur.....	14
b.	Histogrammes.....	14
c.	Les Moments statistiques.....	14
1.1.2.	Descripteurs de textures	14
a.	Filtres de Gabor	15
b.	Histogramme de Gradient Orienté HOG	15
1.1.3.	Descripteurs de Formes.....	16
1.2.	Caractéristiques locale	16
1.2.1.	Les descripteurs des points d’intérêts.....	17
a.	Descripteur SIFT.....	17
b.	Descripteur SURF.....	17
1.3.	Le Vecteur descripteur.....	18
2.	Représentation des caractéristiques d’images	19
2.1.	Modèle de sac de mots visuels (BOW)	19
2.2.	Vecteur de descripteurs agrégés localement (VLAD)	20
3.	Apprentissage supervisé	21
3.1.	Machine à vecteurs de support (SVM)	22
3.2.	K plus proche voisin (KNN)	22
	Conclusion	23

Chapitre III: Conception

Introduction.....	25
1. Architecture générale proposée	25
1.1. Phases d'apprentissages.....	26
1.2. Phases de reconnaissance	31
Conclusion.....	31

Chapitre VI: Implémentation et réalisation

Introduction.....	33
1. Environnement du développement.....	33
2. Base d'images et zone d'étude	33
2.1. Base d'images.....	33
2.2. Zone d'étude	33
3. Résultats obtenues.....	34
3.1. La précision.....	34
3.2. Comparaison des résultats	34
4. Présentation d'application.....	35
Conclusion.....	38
Conclusion générale & perspectives.....	39
Bibliographie	41

Liste des figures

I. Chapitre I :

Figure I.1:principe de base de la télédétection	02
Figure I.2:Exemple de structure 'non dommage' utilisée dans l'étude à trois résolutions spatiales différents (a) 0.35m, (b) 2.0m et (c) 4.0m.....	03

II. Chapitre II:

Figure II.1:l'image (a) représente un descripteur global, et l'image(b) représente un descripteur local	13
Figure II.2:Exemple d'histogrammes pour deux images.....	14
Figure II.3: Etapes du descripteur HOG	16
Figure II.4: Etapes de la méthode SIFT	17
Figure II.5: Etapes de la méthode SURF.....	18
Figure II.6: Généralisation de vocabulaire	20
Figure II.7: Généralisation d'histogramme de chaque image	20
Figure II.8: Images et descripteurs VLAD correspondants, fork=16 centroids (D = 16 x 128).Les composants du descripteur sont représentés comme SIFT, avec des composants négatifs en rouge.....	21
Figure II.9 : Séparateur linéaire entre deux ensembles: les carrés et les points	22

III. Chapitre III :

Figure III.1:Architecture globale de système proposé.....	25
Figure III.2:Exemple points d'intérêt SURF.....	27
Figure III.3:Les Points d'intérêt en vert; sous-région 4X4 en rouge	28
Figure III.4:Formation du HOG d'une image: (a) découpage en cellules et formation de l'histogramme par bloc, (b) vote d'un pixel selon l'orientation de ses gradients.....	29
Figure III.5:L'algorithme K-Means.....	30
Figure III.6:Représentation en histogramme du vocabulaire visuel	30

IV. Chapitre IV:

Figure IV.1: MATLAB.....	33
Figure IV.2: l'interface de démarrage de notre système	36
Figure IV.3: Interface GUI_SURF.....	36
Figure IV.4: Affichage des résultats de classification (SURF)	37
Figure IV.5: Interface GUI_HOG.....	37
Figure IV.6: Affichage des résultats de classification (HOG)	38

Liste des Tableaux

I. Chapitre I :

Tableau I. 1:Classification des dégâts selon l'Échelle Macrosismique Européenne (EMS-98) Pour les bâtiments en maçonnerie (à gauche) et en béton armé (à droite).....	09
---	----

II. Chapitre IV :

Tableau IV.1:Variation de la précision de classification	35
--	----

Liste d'abréviations

HOG: Histogramme De Gradient Orienté.

SIFT: Scale Invariant Feature Transform.

SURF: Speeded Up Robust Features.

BOVW: Bag Of Visuel Word.

VLAD: Vector Of Locally Aggregated Descriptors.

SVM: Support Vector Machines.

Introduction générale

L'extraction c'est une représentation de l'image sous forme de descripteurs et le calcul de leur similarité offre un grand avantage à la recherche d'image par le contenu et à la classification de ces images, beaucoup de domaines peuvent bénéficier de ces derniers, comme : la classification d'image dans une catégorie déterminée, et la recherche dans les bases d'image, la reconnaissance et extraction des formes et des objets, la télédétection des zones pour les études dans plusieurs domaines...etc.

A cet effet, les notions de recherche par le contenu, de classification et d'annotation des images constituent les éléments clés de toute solution voulant répondre aux attentes des professionnels exerçant dans des domaines d'application de l'analyse des images. L'analyse des images en général et des images télédétection en particulier, pour des fins de détection et extraction des objets, est devenue un axe de recherche très important.

La détection automatique des bâtiments endommagés à partir d'images d'télédétection (aériennes et satellitaires) est un problème important pour les planificateurs de sauvetage et le personnel militaire. Une méthode de détection des bâtiments endommagés à partir d'images satellitaires est proposée dans ce mémoire.

Objectifs :

Notre objectif principal de notre travail était le développement d'un système de reconnaissance des dommages causés par les séismes dans les bâtiments en utilisant des caractéristiques bas niveau extraites d'images satellitaires à l'aide d'un classificateur supervisé.

L'objectif secondaire est de comparer la technique d'extraction des caractéristiques locales SURF et globales HOG.

Organisation du mémoire :

Notre mémoire est organisée en quatre chapitres incluant cette Introduction et se termine par une conclusion générale, comme suit :

- Le premier chapitre : il présente quelque notion de base sur l'image l'télédétection et les domaines d'application et son importance dans l'extraction des bâtiments et la détection des dommages.
- Le seconde chapitre : quelques descripteurs (Locaux et globaux) et représentation des caractéristiques de l'image, présente les deux méthodes (SVM, KNN), ainsi que les différentes étapes qui leurs constituent.

- Dans le troisième chapitre, nous l'avons consacré à la conception et l'architecture générale de notre système.
- Enfin, dans le chapitre quatrième: la réalisation de notre application en utilisant l'environnement MATLAB, pour faciliter l'implémentation de la classification des images. Ensuite, nous avons montré les différentes interfaces de cette application. Puis nous avons présenté en images les résultats de nos expérimentations, et nous avons défini un protocole d'évaluation de ces résultats.

En conclusion , nous présentons les points de ce travail et quelques perspectives.

***Chapitre I: Détection de dommage
basé sur les images télédétection***

Introduction :

Les techniques de télédétection jouent un rôle important dans l'obtention d'informations sur les dommages aux bâtiments en raison de leur absence de contact, de leur faible coût, de leur large champ de vision et de leurs capacités de réponse rapide. Maintenant, que plus et divers types de données de télédétection deviennent disponibles, différentes méthodes sont conçues et rapportées pour l'évaluation des dommages à la construction.

Dans ce chapitre, nous nous allons expliquer l'état de l'art dans le domaine de la détection des bâtiments et l'évaluation des dégâts.

1. Technologies de la télédétection :

Le mot 'télédétection' (en anglais « remote sensing ») désigne ensemble des connaissances et techniques utilisées pour déterminer des caractéristiques physiques et biologiques d'objets par des mesures effectuées à distance, sans contact matériel avec ceux-ci.

Elle peut être effectuée depuis une plateforme aérienne (ballon ou avion), depuis un satellite, ou depuis n'importe quel type d'instrument permettant l'acquisition d'informations sur l'environnement. [01]

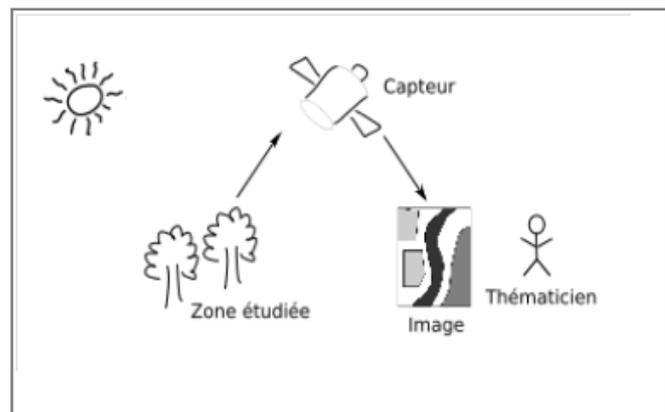


Figure I.1 – principe de base de la télédétection.

2. Résolutions :

La résolution des images numériques est définie par le nombre de pixels par millimètres. En télédétection, on exprime la résolution des images par la taille de la zone couverte par un pixel. Chaque pixel de l'image correspond à une partie de la surface de la terre. On parle alors de "résolution-sol".[02]

Les caractéristiques fondamentales des images de télédétection sont :

La résolution spectrale et la résolution spatiale, la résolution temporelle, métadonnées et règles relatives...etc.

2.1. La résolution spatiale :

La résolution spatiale d'un capteur correspond à la taille minimum des objets pouvant être identifiés au sol, ce qui correspond en télédétection à la taille des pixels d'une image satellite (Picture élément). La capacité à distinguer deux objets voisins différents dépend donc de la résolution spatiale du capteur.

Celle-ci peut varier du kilomètre, pour des capteurs à basse résolution spatiale (tels que NOAA AVHRR, SPOT Vegetation...), utilisés le plus souvent pour des applications météorologiques ou pour cartographier la couverture terrestre à l'échelle globale, au mètre pour des capteurs à haute et très haute résolution spatiale (HRS et THRS).[03]

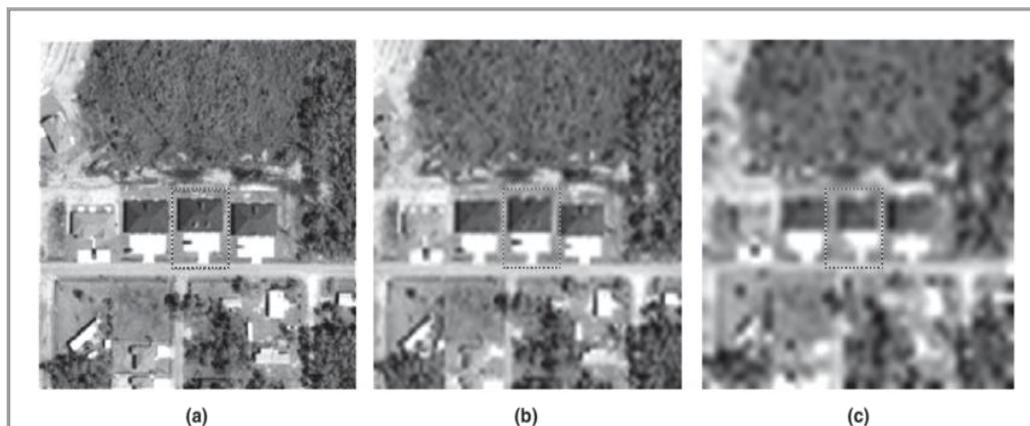


Figure I.2: Exemple de structure "Non dommage" utilisée dans l'étude à trois résolutions spatiales différentes: (a) 0,35 m, (b) 2,0 m et (c) 4,0 m.

2.2. La résolution spectrale :

La résolution spectrale est la capacité d'un capteur à distinguer deux longueurs d'onde voisines [des rayonnements électromagnétiques de fréquences différentes.]. Plus le capteur est sensible à des différences spectrales fines (intervalles de longueur d'onde étroits), plus la résolution spectrale du capteur est élevée. La résolution spectrale est d'autant plus élevée que le nombre de canaux ou bandes spectrales est important. Ainsi, un capteur panchromatique contient un seul canal à large spectre et possède donc une faible résolution spectrale. [03]

La résolution spectrale dépend du dispositif de filtrage optique qui décompose l'énergie captée en bandes spectrales plus ou moins larges.

2.3. La résolution temporelle :

En plus de la résolution spatiale, spectrale, l'autre concept important en télédétection est celui de la résolution temporelle.

La résolution temporelle ou répétitivité, qui correspond à la capacité de revisite d'un capteur sur un même site d'étude, est inversement proportionnelle à la résolution spatiale. Elle est définie

par la répétitivité du positionnement du capteur sur un même lieu à une même heure. Dans le cas des capteurs spatiaux de télédétection passive, on distingue la répétitivité temporelle réelle correspondant à la période de revisite d'un même lieu et la répétitivité temporelle effective dépendant des facteurs climatiques (en présence d'un couvert nuageux, l'information enregistrée par les capteurs n'est pas exploitable). En ce qui concerne les capteurs aéroportés, les missions d'acquisition dépendent, elles aussi, des conditions climatiques mais elles ne répondent à aucune règle de temporalité. En conséquence, les programmations d'acquisitions aéroportées sont plus souples.[03]

3. Acquisition des données de télédétection :

Les moyens d'acquisition d'images de télédétection sont nombreux et performants. On recense pour cela des moyens optiques, lasers, radars, sismographes, gravimètres, etc..

3.1. Les images satellites :

Les images satellites à très haute résolution spatiale (THRS) ont commencé à être commercialisées à partir des années 2000. Leur principale caractéristique est leur résolution spatiale métrique ou Sub-métrique. Elles sont offertes généralement avec une bande panchromatique et quatre bandes spectrales en bleu, vert, rouge et infrarouge.[02]

3.2. Les photographies aériennes :

Avant l'avènement des capteurs spatiaux, les photographies aériennes constituaient l'essentiel des données à THRS. Elles permettent d'enregistrer les entités anthropiques et naturelles en constante évolution qui se trouvent à la surface de la terre.

Les photographies aériennes mettent en évidence les entités comme les montagnes, les canyons et les basses plaines, les cours d'eau, de la source à l'embouchure, elles révèlent les ressources terrestres, comme les lacs, les forêts et la végétation, elles permettent enfin de reconnaître les densités de population. Elles sont aujourd'hui indispensables pour retracer finement l'historique de l'occupation et de l'utilisation des surfaces terrestres avec un niveau de détail élevé ; elles continuent à fournir des informations complémentaires à celles des images satellites ; en outre, elles restent très intéressantes pour l'observation sur des petites surfaces, car les missions d'acquisition peuvent être réalisées avec l'utilisation de plateformes moins coûteuses que les avions telles que des drones.

Les photographies aériennes sont acquises principalement à partir d'avions. Néanmoins, ce terme inclut toute autre photographie qui a fait l'objet d'une acquisition à partir d'une plateforme aéroportée (drone, ULM, ballon, cerf-volant...).[02]

3.3. Comparaison entre photographies aériennes et images satellites :

La résolution d'une photographie ou d'une image nous indique la qualité (du détail visible) sur cette photographie ou image. On mesure cette résolution par la taille minimale que doit avoir un détail observé pour former un point distinct sur l'image.

En photographie aérienne, les prises de vue se font généralement par avion, c'est-à-dire à une altitude pouvant aller jusqu'à 10 Km. La résolution sera bonne et la photographie pourra comporter beaucoup de détails. A contrario, un satellite se trouve beaucoup plus éloigné de la surface du sol. Le satellite SPOT 3 par exemple, se trouve à une altitude de 822 Km. Sa résolution maximale est de 20 mètres ce qui signifie que tout ce qui est inférieur à cette taille sera invisible. Contrairement aux capteurs optiques aéroportés, les capteurs spatiaux peuvent faire des acquisitions en déporté, ce qui leur permet d'avoir une plus large résolution temporelle. Néanmoins, sans compter les difficultés liées aux corrections géométriques nécessaires, notons qu'une déportation trop importante par rapport à une acquisition réalisée directement sous le capteur entraîne des effets de perspectives non négligeables (ombres, côtés des bâtiments, ...).

Enfin, la présence d'un couvert nuageux peut aussi retarder l'acquisition (d'une journée à deux semaines) selon les caractéristiques d'orbite du satellite. Même s'il est possible de programmer des périodes d'acquisition, l'utilisateur maîtrise moins les conditions d'acquisition d'images satellites que, celles liées à l'acquisition de photographies aériennes.[02]

4. Domaine d'applications :

La télédétection est un domaine de recherche récent en pleine expansion qui couvre une large palette d'applications. La littérature est riche en travaux de recherche portant sur la détection du bâti, des réseaux routiers, des zones de végétation, de l'analyse de dégâts, etc.

4.1. Extraction des bâtiments :

L'analyse des images de télédétection en vue de la détection/extraction de bâtiments en milieu urbain, constitue un champ d'investigation important et en pleine expansion. De nombreuses applications comme l'urbanisme ou la cartographie poussent les recherches liées à ce domaine.

Les approches de détection de bâtiments à partir d'imagerie aérienne et satellites peuvent être automatiques ou supervisées. Elles sont basées sur l'extraction de caractéristiques géométriques (comme des lignes ou des coins, ou sur des approches régions). Nous trouvons aussi des méthodes dites hybrides qui utilisent par exemple à la fois des caractéristiques géométriques et des régions ou encore combinant des phases automatiques et semi-automatiques ou supervisées.[02]

4.2. Détection des dommages :

L'imagerie de télédétection est un outil de grand intérêt permettant une estimation rapide des dégâts et une observation détaillée sur une large couverture. A titre d'exemple, un certain nombre de travaux traitent de la détection des dégâts sur les bâtiments, sur les réseaux routiers ou sur l'analyse des risques d'inondation. En effet, plusieurs avantages considérables liés à l'exploitation de l'imagerie de télédétection dans ces domaines peuvent être soulignés comme la possibilité d'étudier les conséquences d'une catastrophe dans des lieux devenus difficiles d'accès par voie terrestre et où les moyens de communication sont souvent bloqués. Par exemple, les tremblements de terre désorganisent les réseaux de transport et interrompent les activités d'urgence comme les secours à victimes, la reconstruction, et les missions de remise en état. Par conséquent, la détection des dégâts du réseau routier postérieurs au tremblement de terre et l'évaluation du niveau de ces dégâts sont deux éléments cruciaux lorsqu'il faut trouver les meilleurs itinéraires et coordonner des missions de sauvetage. Les données de télédétection présentent également des avantages pour l'observation de destructions en zones non sécurisées.[02]

4.3. Extraction des réseaux routiers :

L'extraction des réseaux routiers a fait l'objet de nombreux travaux en traitement d'images et en télédétection. Les réseaux routiers apportent des informations contextuelles de première importance, permettant de localiser d'autres types d'éléments sur une image de télédétection, telles que les structures domiciliaires ou commerciales. Même aujourd'hui, la cartographie du réseau de rues est essentiellement effectuée par interprétation d'images avec un apport de mesure de terrain. Cette procédure est un processus cher et gourmand en temps et demande à être ré-effectué à chaque mise à jour du réseau.[02]

L'intérêt pour une méthode automatique et robuste d'extraction des routes à partir d'images de télédétection est indéniable. Une grande variété des méthodes d'extraction des réseaux routiers à partir d'images aériennes et satellites a vu le jour ces dernières années.

4.4. La modernisation de l'agriculture :

L'imagerie de télédétection est aussi un outil déterminant pour la modernisation de l'agriculture. En effet, la télédétection peut fournir des données très utiles pour identifier, surveiller les zones cultivées et prendre des décisions sur les cultures et les stratégies agricoles. Le principal défi est alors de répondre aux problèmes de l'identification des carences en éléments nutritifs, de maladies, de manque ou d'excédent d'eau, d'invasion de mauvaises herbes, de dégâts d'insectes, dus au vent ou aux herbicides, etc.[02]

En plus des domaines d'application mentionnés peuvent des applications liées à la cartographie du littoral (la ligne de rivage) et à la détection des changements intervenant sur ce dernier. L'extraction du littoral est essentielle pour plusieurs applications à savoir la navigation autonome de véhicules aériens, le géo-référencement des images, la surveillance de la géomorphologie côtière, la protection de l'environnement côtier, etc.

5. Détection et évaluation des dommages après le séisme:

Méthodes actuelles de détection et d'évaluation des dégâts que ce soit par une analyse sur le terrain par des équipes de relevés ou par une interprétation visuelle sur des images satellitaires ou acquises par des capteurs aéroportés, l'évaluation des dégâts est accomplie en se basant sur une ou plusieurs échelles de destruction.

5.1. Recherche sur le terrain :

Le moyen le plus direct est d'envoyer une ou plusieurs équipes directement sur les lieux de la catastrophe. Les équipes peuvent évaluer l'intégrité des structures, noter les dégâts apparents de l'extérieur et dresser un bilan de la résistance des bâtiments aux secousses sismiques selon les types de matériaux utilisés. De prime abord, il s'agit d'une opération dispendieuse et surtout très longue. De plus, l'accès à certaines zones peut être impossible en raison de restrictions d'ordre politiques ou pratiques.

5.2. Détection de dommage par les images de télédétection :

La technologie de télédétection est la technologie la plus efficace pour les agences de secours en cas de catastrophe et les unités de protection civile à évaluer les dommages à des fins de planification.

Surtout à la suite de séismes extrêmes, en termes de fournir des informations critiques pour les intervenants d'urgence, La quantification rapide de l'étendue et de la gravité des dommages au bâtiment est une priorité élevée.

Une option plus intéressante et de nos jours beaucoup plus accessible est l'analyse visuelle des dégâts sur des images acquises par des systèmes d'observation de la terre. Il existe organismes travaillant à l'analyse visuelle des images, ces organismes doivent agir rapidement et offrir une analyse juste de la zone affectée. Ils font malheureusement face à plusieurs problèmes lorsqu'une catastrophe survient. Il y a un délai parfois de plusieurs jours entre la programmation du satellite, l'acquisition des images de la zone désirée et la réception des données.[04]

a. Télédétection et dommage grades :

Il existe plusieurs possibilités pour évaluer les dégâts sur les bâtiments comme la méthode binaire (intact ou endommagé) ou la méthode à degrés (intact, endommagé, détruit), Mais dans les images de télédétection sont détectées des dommages et évalués soit intact ou complètement endommagé.

b. Dommage grades :

Pour caractériser les dommages il est nécessaire d'établir une typologie de la gravité des dégâts aux bâtiments en définissant une échelle de référence. Au-delà d'une simple distinction entre bâtiments sains et bâtiments détruits, de nombreuses études S'appuient sur l'échelle définie par l'échelle macrosismique européenne 1998 (l'EMS-98) (Tableau. I.1) qui distingue 5 degrés d'endommagement (légère dégâts, dégâts modérés, dégâts importants, dégâts très importants, et destruction). Cette gradation des dégâts dus aux séismes via l'EMS-98 est intéressante puisqu'elle représente une homogénéisation internationale de la nomenclature des degrés de destruction. Cependant les études signalent que la distinction en 5 degrés même à partir d'images THR est impossible, les niveaux 1 et 2 sont confondus, de même pour les niveaux 3 et 4. Ces confusions proviennent du fait que certains dégâts mineurs (fissures ou dégâts non structurels) ne sont pas visibles, de même pour les dégâts au niveau des façades qui ne sont visibles qu'à partir d'images obliques. De plus les bâtiments effondrés de façon symétrique (effondrement d'un ou plusieurs étages appelés "étages mous") ne sont pas toujours identifiables, le toit conservant la même forme. En revanche les niveaux 4 et 5 sont généralement bien identifiables. [05]

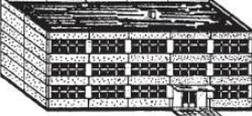
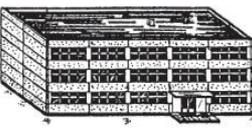
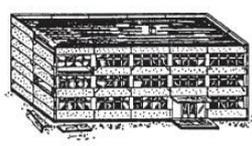
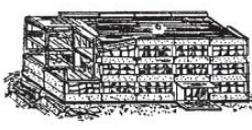
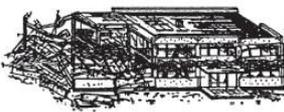
	<p>Degré 1: Dégâts négligeables à légers (aucun dégât structural, légers dégâts non structuraux)</p> <p>Fissures capillaires dans très peu de murs. Chute de petits débris de plâtre uniquement. Dans de rares cas, chute de pierres descellées provenant des parties supérieures des bâtiments.</p> <p>Fissures fines dans le plâtre sur les parties de l'ossature ou sur les murs à la base. Fissures fines dans les cloisons et les remplissages.</p>	
	<p>Degré 2: Dégâts modérés (dégâts structuraux légers, dégâts non structuraux modérés)</p> <p>Fissures dans de nombreux murs. Chutes de grands morceaux de plâtre. Effondrement partiel des cheminées.</p> <p>Fissures dans les structures de types portiques (poteaux et poutres) et dans structures avec murs. Fissures dans les cloisons et les murs de remplissage; chute des revêtements friables et du plâtre. Chute du mortier aux jonctions entre les panneaux des murs.</p>	
	<p>Degré 3: Dégâts sensibles à importants (dégâts structuraux modérés, dégâts non structuraux importants)</p> <p>Fissures importantes dans la plupart des murs. Les tuiles des toits se détachent. Fractures des cheminées à la jonction avec le toit; défaillance d'éléments non structuraux séparés (cloisons, murs pignons).</p> <p>Fissures dans les poteaux et dans les nœuds à la base de l'ossature et aux extrémités des linteaux des murs avec des ouvertures. Ecaillage du revêtement de béton, flambement des barres d'armature longitudinale.</p> <p>Fissures importantes dans les cloisons et les murs de remplissage, défaillance de certains panneaux de remplissage.</p>	
	<p>Degré 4: Dégâts très importants (dégâts structuraux importants, dégâts non structuraux très importants)</p> <p>Défaillance sérieuse des murs; défaillance structurale partielle des toits et des planchers.</p> <p>Fissures importantes dans les éléments structuraux avec défaillance en compression du béton et rupture des barres à haute adhérence; perte de l'adhérence barres-béton; basculement des poteaux. Ecoulement de quelques poteaux ou d'un étage supérieur.</p>	
	<p>Degré 5: Destruction (dégâts structuraux très importants)</p> <p>Effondrement total ou presque total.</p> <p>Effondrement total du rez-de-chaussée ou de parties de bâtiments.</p>	

Tableau. I.1 – Classification des dégâts selon l'Échelle Macrosismique Européenne (EMS-98) pour les bâtiments en maçonnerie (à gauche) et en béton armé (à droite).

5.3. Méthode de détection de dommage :

La détection de dommages à partir d'images de télédétection peut-être effectuée soit manuellement par photo-interprétation, soit automatiquement à l'aide de différentes approches.

a. Interprétation visuelle :

La photo-interprétation consiste en une analyse visuelle de l'image post-crise par un opérateur (ou photo-interprète) qui assigne un niveau de dégâts à chaque bâtiment. L'opérateur peut utiliser une image pré-crise pour constater des changements. Les indices généralement utilisés sont la linéarité des bords des toits, l'homogénéité de sa texture, la forme des ombres du bâtiment et des éléments tels que les cheminées ou les réservoirs d'eau, la présence de gravats ou encore le mouvement anormal du toit d'une image à l'autre. Le principal problème de la photo-interprétation est qu'elle sous-estime les dégâts. Un autre inconvénient de l'approche manuelle est qu'elle est longue et subjective, les résultats variant d'un interprète à un autre.

Par contre le cerveau humain arrive facilement à reconnaître des objets similaires sur plusieurs images, à les associer et ainsi à constater des changements. Ces concepts qui paraissent simples ne sont pas aussi évidents à mettre en œuvre dans les méthodes automatiques. Ainsi l'approche manuelle reste considérée par les chercheurs comme une méthode fournissant des résultats

fiables, plus facile à appliquer que les méthodes automatiques, bien que plus longue et subjective.[05]

b. Techniques multi-temporelles :

Les techniques dites multi-temporelles s'appuient sur des méthodes de détection de changements. Elles consistent à détecter les différences des caractéristiques spectrales d'un même objet situé dans des images acquises avant et après la crise (la comparaison d'images avant et après crise).

Les techniques proposées peuvent être classées en deux catégories. La première catégorie regroupe des techniques généralement rapides et qui consistent à comparer successivement chaque pixel de l'image avant-crise avec le pixel correspondant dans l'image après-crise. Ces techniques présentent l'inconvénient d'être sensibles aux changements non pertinents (un dégât est souvent caractérisé par un changement de radiométrie excédant un seuil défini manuellement ou automatiquement), ce qui engendre de nombreuses fausses alarmes sur les résultats obtenus. La deuxième catégorie regroupe les techniques dites orientées-objets qui consistent à découper l'image en zones homogènes par segmentation selon certains critères faisant appel généralement à des grandeurs radiométriques. Dans ce cas, l'analyse de degré d'un dégât est basée sur des caractéristiques de types radiométrique et géométrique. Ces méthodes sont moins sensibles aux changements non pertinents car elles tiennent en compte de la distribution spatiale des pixels. Néanmoins, elles sont sensibles aux résultats de segmentation et à la mise en correspondance des objets extraits.[01]

c. Techniques mono-temporelles (uni-temporelle):

L'analyse mono-temporelle n'emploie pas d'image de référence pré-crise, et se base sur une(ou plusieurs) image(s) de crise. Ce type d'analyse présente l'avantage de ne nécessiter aucune image de référence Parce que dans certains cas, ils ne sont pas disponibles pour certaines zones du globe. Les approches développées dans ce sens sont généralement destinées à détecter des zones brûlées ou à quantifier des dégâts sur le bâti suite à un séisme. L'analyse de l'image post-crise en vue de détecter les zones brûlées est fondée sur les différences des signatures spectrales des différents objets présents dans l'image. La détection et la quantification des dégâts sur le bâti s'appuient sur différents critères [01].

En cas de dégâts, les toits de bâtiments présentent d'une part des irrégularités sur leur contour, d'autre part une plus grande hétérogénéité radiométrique par rapport à ceux qui sont intacts. Par ailleurs, l'information sur l'ombre peut être exploitée dans le cas où aucun signe de dégâts n'est

déecté sur le toit. Plus exactement, l'absence ou la taille anormalement petite (en comparaison des bâtiments voisins) de l'ombre issue d'un bâtiment indique que celui-ci s'est peut effondré sur lui-même. Comme pour l'analyse multi-temporelle, deux sortes de catégories de méthodes peuvent être distinguées, à savoir les méthodes travaillant à l'échelle du pixel et les méthodes basées-objets.[01]

Conclusion :

La télédétection est un domaine de recherche récent en pleine expansion qui couvre une large palette d'applications. Ce chapitre décrit de façon les technologies de la télédétection, les méthodes actuelles d'évaluation des dégâts et la détection des bâtiments et l'évaluation des dégâts sur ces derniers lors de catastrophes. Nous avons rappelé le principe général du processus d'acquisition des données de télédétection ainsi que les différents types de données manipulées à savoir les images satellites, les images aériennes ; parce que la télédétection, jouent un rôle majeur dans l'évaluation des dommages causés par les tremblements de terre que ce soit en analyse visuelle ou avec un logiciel de traitement.

Nous présentons dans le chapitre suivant un état de l'art sur les méthodes d'extraction et représentation des caractéristiques d'images.

*Chapitre II: Extraction des
Caractéristiques d'images*

Introduction :

Les contenus visuels des images sont extraits et décrits par des vecteurs multidimensionnels de caractéristiques, et ces vecteurs forment la base de données de descripteurs (ou signatures) l'extraction des contenus visuels des images se concentre généralement sur les caractéristiques de bas niveau. Il y a deux approches principales de caractériser : la première est la construction de descripteurs globaux à toute l'image et la seconde est locale et consiste à calculer des caractéristiques sur des portions restreintes de l'image et nous avons expliqué en détail dans ce chapitre.

1. Extraction des caractéristiques :

La description du contenu des images est une étape essentielle dans un système de recherche d'image par le contenu, car la performance de système dépend en grande partie du choix des descripteurs employés et des techniques associées à leurs extractions. Un descripteur est défini comme la connaissance utilisée pour caractériser l'information contenue dans les images.

Cette étape permet de fournir une représentation du contenu de l'image appelé aussi signature de l'image.[06]

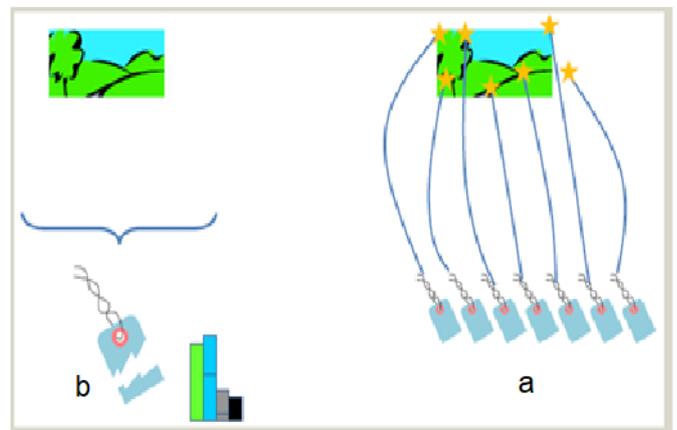


Figure II.1: l'image (a) représente un descripteur global, et l'image(b) représente un descripteur local

Pour décrire l'image, des descripteurs de bas-niveau sont utilisés, appelés aussi vecteurs caractéristiques, tels que la couleur, la texture et la forme et c'est caractéristiques globales et caractéristiques locales :

1.1. Caractéristiques globales :

Ces caractéristiques concernant la recherche sur toute l'image comme la couleur, texture, forme, ...etc. Un système basé uniquement sur des caractéristiques globales ne peut pas donner les résultats désirés.

1.1.1. Descripteurs de la couleur :

La couleur est une caractéristique riche d'information et très importante et très utilisée pour la représentation l'image RS aérienne et autre image.

La couleur est devenue la première signature employée pour la recherche d'images par le contenu en raison de son invariance par rapport à l'échelle, la translation et la rotation [07]. Ces

valeurs tridimensionnelles font que son potentiel discriminatoire soit supérieur à la valeur en niveaux de gris des images. Une indexation couleur repose sur deux principaux choix : l'espace colorimétrique et le mode de représentation de la couleur dans cet espace [08].

a. L'espace de Couleur :

Il existe plusieurs espaces colorimétriques qui ont chacun certaines caractéristiques intéressantes l'espace RGB ; l'espace HSI ...

L'avantage majeur de la fonctionnalité d'espace de couleur est son invariant avec la rotation et l'échelle.

b. Histogrammes :

Le descripteur de couleurs le plus courant et le plus rencontré dans la littérature est l'histogramme de couleurs. L'histogramme est un outil statistique qui permet d'estimer la densité de probabilité d'une distribution à partir d'échantillons. L'intervalle des valeurs possibles est divisé en classes, puis pour chacune d'elles on compte le nombre d'échantillons associés.[09]

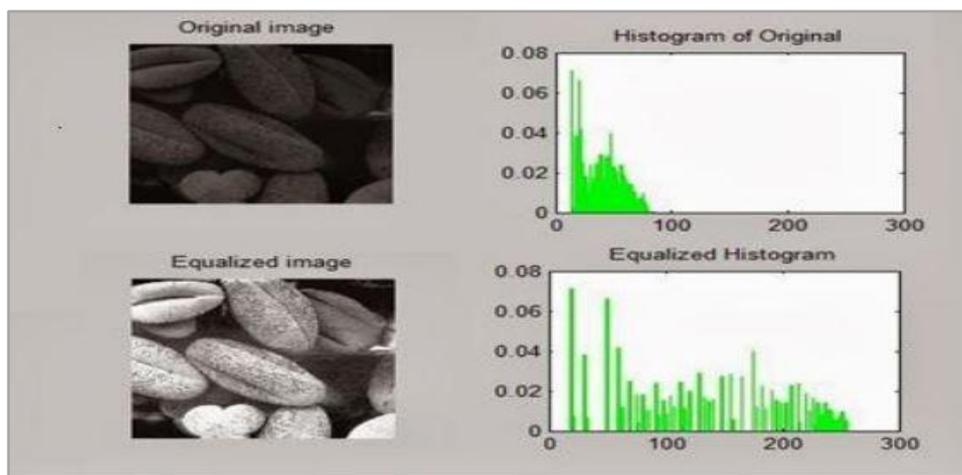


Figure II.2: Exemples d'histogrammes pour deux images.

c. Les Moments Statistiques :

La méthode d'histogramme utilise la distribution complète de la couleur. On doit stocker de nombreuses données. Au lieu de calculer la distribution complète, dans les systèmes de recherche d'images, on calcule seulement des caractéristiques dominantes de couleur tel que l'espérance, la variance et d'autres moments.[09]

1.1.2. Descripteurs de Textures :

La texture est largement utilisée dans le traitement d'image RS, et il est l'une des caractéristiques typiques des différentes entités terrestres. Par rapport aux images générales, la texture de l'image RS est plus complexe et abondante et est plus sensible à différentes entités. La

texture peut être utilisée efficacement en RS description d'image et mesure dans la récupération d'image.

De nombreuses méthodes pour analyser les caractéristiques de texture dans l'image RS ont proposé, y compris la matrice de cooccurrence grise, champ aléatoire Markov, transformation de Gabor, fractale, ondelette, unité de texture et autres. Parmi toutes les méthodes, gris la matrice de co-occurrence, les méthodes fractales et ondelettes sont utilisées plus largement.[11].

a. Filtres de Gabor :

Les filtres de Gabor sont largement utilisés en indexation, pour la description de la texture. Ils permettent une bonne résolution temporelle à haute fréquence et une bonne résolution harmonique sans grande précision temporelle à basse fréquence [10]. Sommairement, les paramètres de texture sont déterminés en calculant la moyenne et l'écart type des niveaux de gris de l'image filtrée par Gabor. En fait, ce n'est pas une seule valeur de moyenne et d'écart type qui sera calculée, mais plutôt un ensemble de valeurs égal au nombre d'échelles multiplié par le nombre d'orientations utilisées. On aura donc ce qui est parfois appelé la banque de filtre de Gabor.

b. Histogramme de Gradient Orienté HOG :

Les HOG ont été proposés par Navneet Dalal et Bill Triggs, chercheurs à l'Inria Grenoble, à la conférence CVPR de juin 2005 [12].

L'histogramme des dégradés orientés est un descripteur de caractéristique utilisé dans la vision par ordinateur et le traitement d'image pour la détection d'objets.[12] l'idée essentielle derrière l'histogramme de gradient orienté c'est que l'apparence locale et la forme d'objet dans une image peut être décrite par la distribution d'intensité des gradients ou de direction des contours. Cette méthode est similaire à celle des histogrammes d'orientation de bord, des descripteurs de transformation d'entités invariables à l'échelle et des contextes de forme, mais diffère en ce qu'elle est calculée sur une grille dense de cellules uniformément espacées. Le descripteur HOG maintient quelques avantages clés par rapport aux autres méthodes. Puisque le descripteur histogramme de gradient orienté opère sur les cellules localisées, la méthode maintient l'invariance à des transformations géométriques et photométriques, ces changements ne feront leur apparition que dans les larges régions d'espaces.[12]

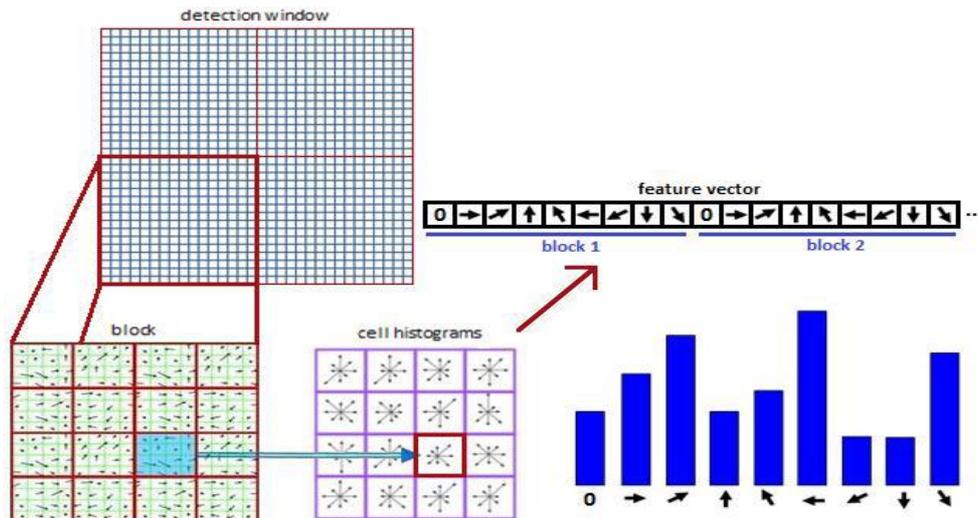


Figure II.3 : Etapes du Descripteur HOG.

1.1.3. Descripteurs de Formes :

La forme est généralement une description très riche d'un objet. L'extraction d'attribut géométrique a été le fer de lance de la recherche d'image par le contenu ces dernières années. De nombreuses solutions ont été proposées pour représenter une forme, nous distinguons deux catégories de descripteurs de formes : les descripteurs basés sur les régions et les descripteurs basés sur les frontières.

Les premiers font classiquement référence aux moments invariants et sont utilisés pour caractériser l'intégralité de la forme d'une région. Ces attributs sont robustes aux transformations géométriques comme la translation, la rotation et le changement d'échelle. La seconde approche fait classiquement référence aux descripteurs de Fourier et porte sur une caractérisation des contours de la forme.[13]

1.2. Caractéristiques local :

La requête partielle basé à la recherche de zones particulières de l'image à l'aide de segmentation en région, détection de points d'intérêt,...etc. Afin d'extraire des caractéristiques locales, il est indispensable de détecter les zones d'intérêt de l'image, se sont les points saillants de l'image, puis de calculer en chacune de ces zones un vecteur caractéristique. Pour la recherche par le contenu, différentes méthodes existent pour définir et extraire des composantes d'images.[06]

1.2.1. Les descripteurs des points d'intérêts :

Les points anguleux (ou points d'intérêt, points saillants,...) Sont des points qui contiennent beaucoup d'information relativement à l'image. Ce sont des points aux voisinages desquels l'image varie significativement dans plusieurs directions.[09]

a. Descripteur SIFT (Scale Invariant Feature Transform):

SIFT est une méthode développée par David Lowe en 2004 , permettant de transformer une image en ensemble de vecteurs de caractéristiques qui sont invariants par transformations géométriques usuelles (homothétie, rotation) et de manière moins fiables aux transformations affines et à l'illumination. Le point fort de la méthode de Lowe est qu'elle est capable de mettre en correspondance des points distants avec des variations de caméra importantes.

L'algorithme des SIFT vient combler en grande partie les limites des méthodes d'extraction de points remarquables déjà développées avant lui par Harris, et plus tard par Mohr et Schmid. En effet, il a contribué à l'amélioration des techniques d'extraction d'information dans une image en apportant un algorithme robuste et satisfaisant les propriétés que requièrent les procédés de vision artificielle notamment le recalage d'images," technique consistant à trouver une transformation géométrique permettant de passer d'une image (dite source) à une autre image (dite cible) " la mise en place de la méthode nécessite deux étapes principales.

Premièrement, il est nécessaire d'extraire les caractéristiques d'un objet et de calculer ses descripteurs, c'est-à-dire, de détecter les caractéristiques qui sont les plus susceptibles de représenter cet objet, de le définir et de le discriminer par rapport aux autres. Deuxièmement, il faut mettre en place une procédure de mise en correspondance (« matching »). C'est le but ultime de la méthode.

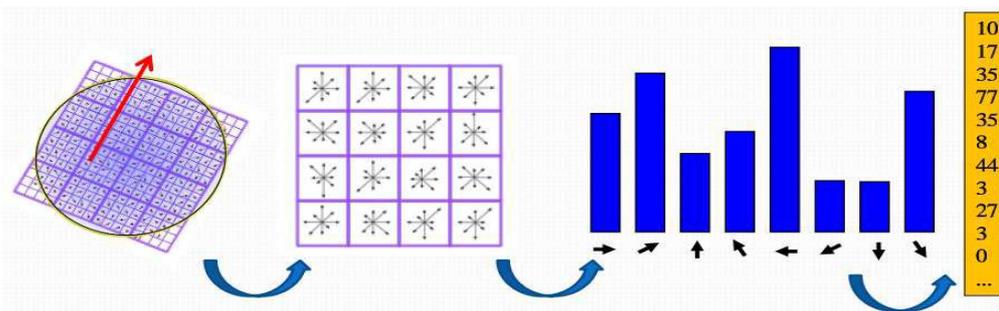


Figure II.4 : Etapes de la méthode SIFT

b. Descripteur SURF (Speeded Up Robust Features):

SURF est un algorithme de détection de caractéristique et un descripteur, présenté par des chercheurs de l'Eth Zurich et de la Katholieke Universiteit Leuven pour la première fois en 2006,

puis dans une version révisée en 2008. Il est utilisé dans le domaine de vision par ordinateur, pour des tâches de détection d'objet ou de reconstruction 3D.[14]

SURF est partiellement inspiré par le descripteur SIFT, qu'il surpasse en rapidité et selon ses auteurs, plus robuste pour différentes transformations d'images. SURF est fondé sur des sommes de réponses d'ondelettes de Haar 2D et utilise efficacement les images intégrales. En tant que caractéristique de base, SURF utilise une approximation d'ondelettes de Haar du détecteur de blob à base de déterminant hessien.

Présenté comme une solution de hautes performances capable de d'approcher et même de dépasser les schémas précédents du point de vue répétitivité, distinctivité et robustesse, SURF (ou Speeded Up Robust Features) est un détecteur de points d'intérêts invariant aux changements d'échelle et aux rotations. Afin de réduire les temps de traitement, il fait usage d'images intégrales pour le calcul de la convolution, et reprend les points forts des meilleurs détecteurs et descripteurs l'ayant précédé, en utilisant des mesures à base de matrices hessiennes rapides pour le détecteur et un descripteur basé sur les distributions. Il décrit une distribution de réponses d'ondelettes de Haar dans le voisinage du point d'intérêt. Seulement 64 dimensions sont utilisées, ce qui permet de réduire le temps de calcul des caractéristiques et de recherche des correspondances. L'étape d'indexation est basée sur le signe du Laplacien, d'où l'accélération de la recherche de correspondances et l'amélioration de la robustesse du descripteur. [14]

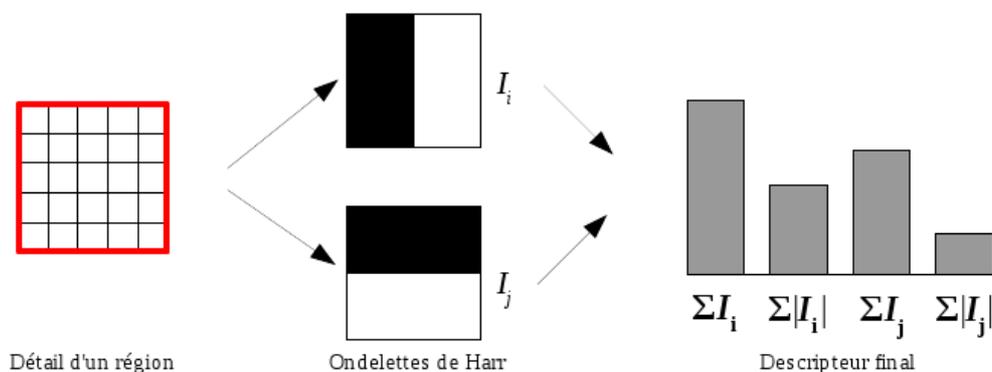


Figure II.5 : Etapes de la méthode SURF.

1.3. Le Vecteur descripteur :

Le vecteur descripteur contient les attributs intéressants extraits de l'image. Il se présente en général sous la forme d'un vecteur à n composantes réelles. Les attributs extraits des images sont de différents types et sont exprimés dans des unités différentes selon qu'ils appartiennent à la couleur, la texture, la forme. Une étape de normalisation est indispensable, elle va permettre de réajuster les valeurs des attributs pour les rendre commensurables.[15]

Dans certaines applications, il n'est pas suffisant d'extraire un seul type de caractéristique pour obtenir les informations pertinentes à partir des données d'image. Au lieu de cela, deux entités

différentes ou plus sont extraites, ce qui donne deux ou plusieurs descripteurs de caractéristiques à chaque point de l'image. Une pratique courante consiste à organiser les informations fournies par tous ces descripteurs en tant qu'éléments d'un seul vecteur, communément appelé vecteur de caractéristiques. L'ensemble de tous les vecteurs caractéristiques possibles constitue un espace caractéristique. un exemple courant de vecteurs de caractéristiques apparaît lorsque chaque point d'image doit être classé comme appartenant à une classe spécifique. En supposant que chaque point d'image possède un vecteur de caractéristiques correspondant basé sur un ensemble approprié de caractéristiques, signifiant que chaque classe est bien séparée dans l'espace caractéristique correspondant, la classification de chaque point d'image peut être effectuée en utilisant une méthode de classification standard.[15]

2. Représentation des caractéristiques d'images :

2.1. Modèle de sac de mots visuels (BOW) :

Le modèle de mots visuels décrit une image en utilisant un ensemble de mots visuels appelés vocabulaire visuel. Le vocabulaire est obtenu en regroupant des entités locales extraites d'images où chaque groupe résultant est un mot visuel. Dans ce modèle, une image est finalement représentée par un histogramme, où chaque bag correspond à un mot visuel et le poids associé représente son importance dans l'image.[16] Ainsi, la construction de l'histogramme nécessite trois étapes:

- 1) extraire les caractéristiques visuelles,
- 2) construire un vocabulaire visuel.
- 3) indexer les images.

Extraction des caractéristiques visuelles :

De nombreuses approches ont été proposées pour extraire des caractéristiques locales à partir d'images. Approche intéressante consiste à extraire les points-clés. Ces points clés sont les centres des taches saillantes puisqu'ils sont généralement situés autour des coins et des bords.

Construire le vocabulaire visuel :

Construire le vocabulaire visuel signifie quantifier les descripteurs locaux extraits pour un large échantillon d'images. Le vocabulaire peut être généré en regroupant les entités SIFT à l'aide de l'algorithme k-means standard. La taille du vocabulaire est le nombre de clusters et les centres de clusters sont les mots visuels. Chaque image dans la base de données sera représentée par des mots visuels de ce vocabulaire.

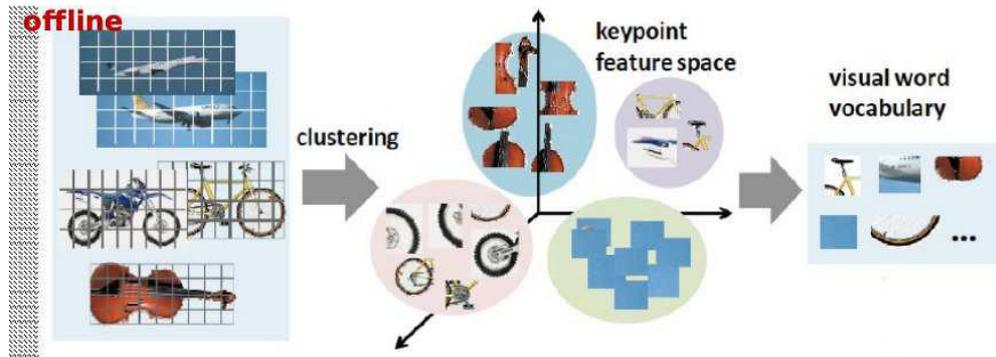


Figure II.6: Génération du vocabulaire

Indexation des images :

Une fois le vocabulaire visuel construit, nous indexons les images en construisant leurs signatures BOVW. Une signature image BOVW nécessite de trouver le poids de chaque mot visuel du vocabulaire. Ainsi, chaque image est représentée par un histogramme où les cases sont des mots visuels et les poids sont leurs fréquences dans l'image.

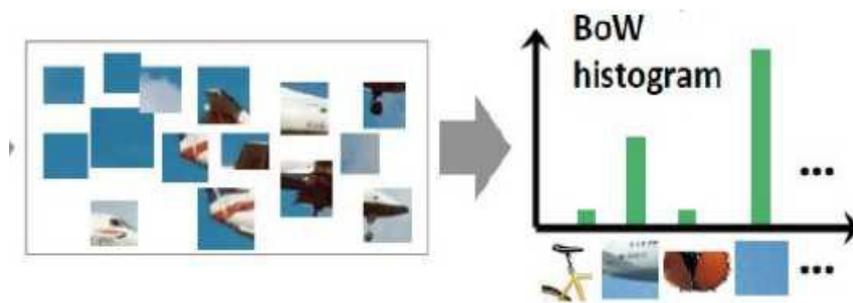


Figure II.7: Génération des histogrammes de chaque image

2.2. Vecteur de descripteurs agrégés localement (VLAD) :

VLAD (vecteur de descripteurs agrégés localement) un descripteur dérivé à la fois du noyau BOF et du noyau de Fisher, Proposé par Jégou et al pour agrèr des descripteurs basés sur un critère de localité dans l'espace caractéristique et produit une représentation compacte.[19]

VLAD est construit comme suit: les régions sont extraites d'une image en utilisant un détecteur invariant affine, et décrites en utilisant le descripteur SIFT 128-D. Chaque descripteur est ensuite affecté au groupe le plus proche d'un vocabulaire de taille k (où k est généralement 64 ou 256, de sorte que les clusters sont assez grossiers). Pour chacune des k grappes, les résidus (différences entre les descripteurs et les centres de grappes) sont accumulés et les sommes de résidus de k 128-D sont concaténées en un seul descripteur de dimension $k \times 128$; nous nous référons à lui comme le VLAD non normalisé.[17]

Il diffère du descripteur d'image BOVW en enregistrant la différence par rapport au centre du cluster plutôt que par le nombre de SIFT affectés au cluster. Il hérite de certaines des invariances

du descripteur SIFT d'origine, telles que l'invariance rotationnelle dans le plan, et est quelque peu tolérant à d'autres transformations telles que la mise à l'échelle de l'image et l'écrêtage. Une autre différence par rapport à l'approche standard BOVW est que les systèmes de récupération VLAD excluent généralement l'utilisation des descripteurs locaux originaux. Ils sont utilisés dans les systèmes BOVW pour la vérification spatiale et le reclassement, mais nécessitent trop de stockage pour être conservés en mémoire sur une seule machine pour de très grands jeux de données d'images. VLAD est similaire dans l'esprit des vecteurs Fisher, comme les deux aspects record de la distribution SIFT affectés à un centre de cluster.[18;19]

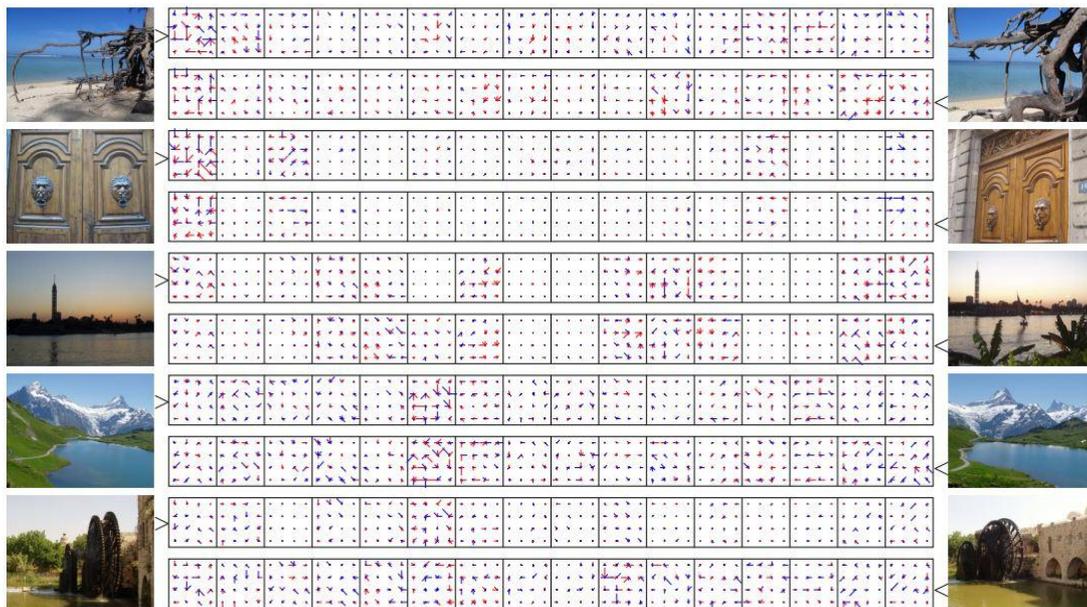


Figure II.8 : Images et descripteurs VLAD correspondants, fork=16 centroids (D = 16 x 128). Les composants du descripteur sont représentés comme SIFT, avec des composants négatifs en rouge.

3. Apprentissage supervisé :

A la différence de l'apprentissage non supervisé, dans le cas de l'apprentissage supervisé l'ensemble des données est divisé en 2 parties:

Données d'apprentissage: un apprentissage a priori est effectué sur ces données avec une intervention de l'humain.

Données de test: utilisées pour vérifier la validité d'apprentissage. L'objectif de l'apprentissage à partir d'exemples étiquetés appelé aussi apprentissage supervisé est de construire une fonction qui permet d'approcher au mieux une fonction inconnue qui génère des données aléatoires, indépendantes et identiquement distribuées et dont nous ne disposons que de quelques exemples. Le problème d'apprentissage à partir d'exemples apparaît dans plusieurs domaines divers et

variés, comme la prédiction, la classification, la reconnaissance de formes, la reconnaissance d'objets, ... etc.

3.1. Machine à vecteurs de support (SVM) :

Les SVM constituent une classe d'algorithmes basée sur le principe de minimisation de « Risque structurel » décrit par la théorie de l'apprentissage statistique qui utilise la Séparation linéaire. Cela consiste à séparer par un hyperplan des individus représentés dans un espace de dimension égale au nombre de caractéristiques, les individus étant alors séparés en deux classes. Cela est possible quand les données à classer sont linéairement séparables. Dans le cas contraire, les données seront projetées sur un espace de plus grande dimension afin qu'elles deviennent linéairement séparables.

Dans l'exemple de la figure. II.9, initialement, les données ne sont pas classées. Après l'exécution du SVM, l'ensemble des données est divisé en deux classes:

Les points et les carres. L'efficacité d'un SVM réside dans la maximisation de la marge (la variable w), à la différence d'un réseau de neurones ou le but est de trouver la ligne séparatrice entre les deux ensembles (ici la ligne noire), un SVM maximise la marge qui est la distance entre la ligne verte et la ligne rouge Sans pour autant augmenter le taux d'erreur.[20]

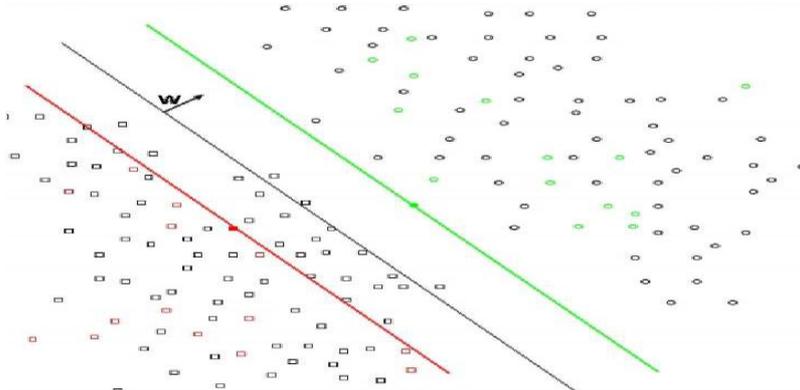


Figure II.9: Séparateur linéaire entre deux ensembles: les carres et les points.

3.2. K plus proche voisin (KNN) :

L'algorithme de k plus proches voisins (KPPV) (K-Nearest-Neighbors (KNN) en Anglais) est une méthode non paramétrique et supervisée de classification. Elle est largement utilisée en classification d'une manière générale et en segmentation d'images en particulier. Elle repose sur un principe simple et intuitif de regroupement d'individus en fonction de leur voisinage. L'algorithme de K Plus Proche Voisin se base essentiellement sur les deux éléments principaux suivants :

1. Le nombre de cas les plus proches (K) à utiliser et une métrique pour mesurer le plus proche voisin.
2. La valeur de K est spécifiée à chaque utilisation de l'algorithme puisqu'il détermine le nombre de cas existants qui sont considérés pour prédire un nouveau cas.

Conclusion :

Dans ce chapitre, nous présentons la représentation d'image et extraction et la description des caractéristiques nécessaires pour comprendre les approches que nous avons utilisée dans ce mémoire. Enfin, nous avons présenté les approches d'apprentissage et les différentes méthodes pour la classification.

Le prochain chapitre, donne les détails de la conception, ainsi que la méthode pour la réalisation de notre application.

Chapitre III: Conception

Introduction :

Dans ce chapitre, nous nous évoquons la conception du système proposé dans ce mémoire à réaliser à un système de classification des images satellites se basant sur la technique SVM et fournir un aperçu plus détaillé de l'architecture générale et les techniques utilisées dans notre travail.

1. Architecture générale proposée:

La Figure III.1 présente les différentes étapes par lesquelles nous sommes passés pour la réalisation de notre système. Nous exposons en détail les différentes étapes de deux phases dans ce qui suit :

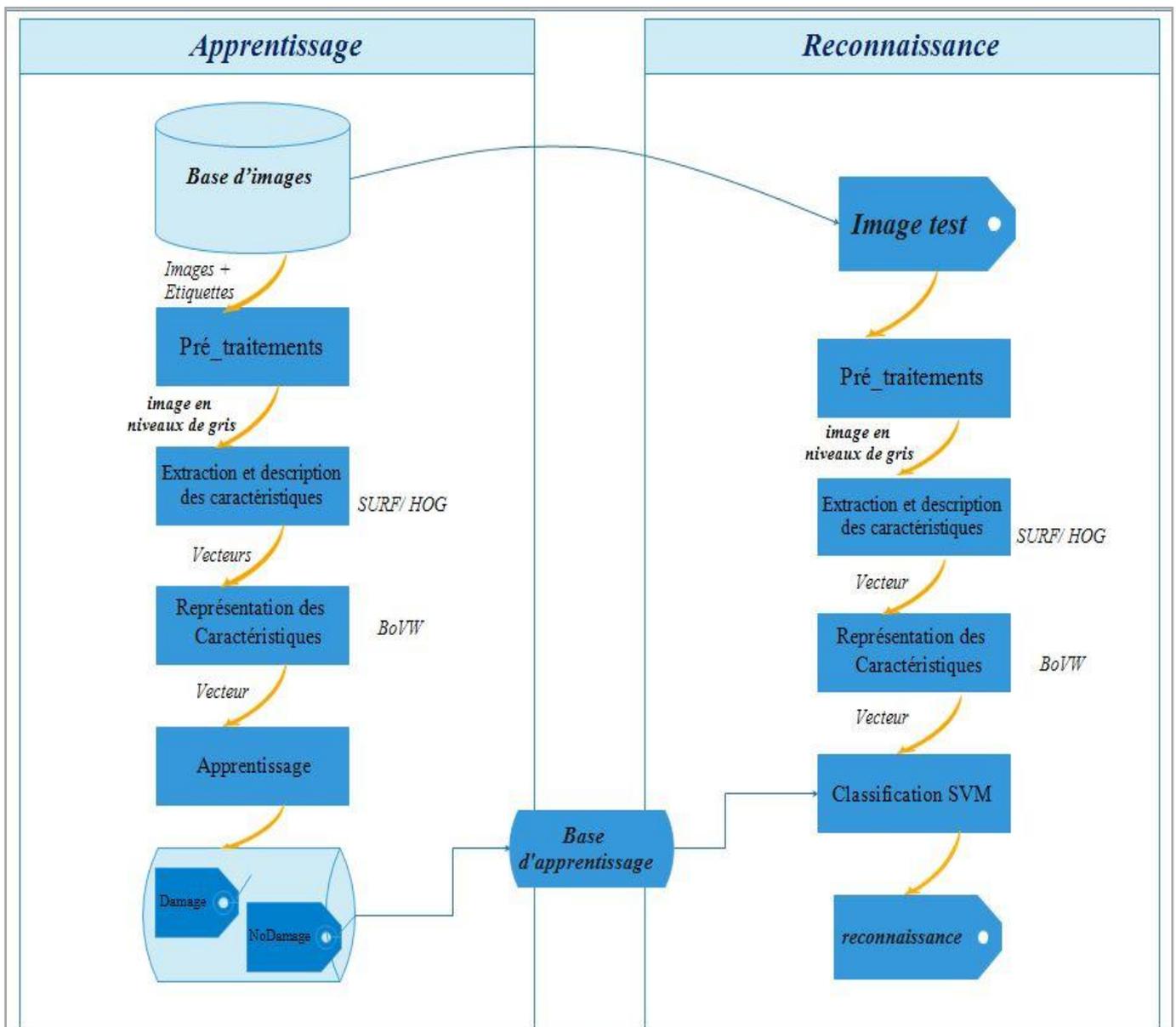


Figure III.1 : Architecture générale de notre système.

1.1. Phases d'apprentissage :

Au cours de phases d'apprentissage, le système proposé est le suivant:

1.1.1. Prétraitement :

Dans cette étape, pour chaque image de l'ensemble de données:

Nous allons redimensionner l'image à la taille standard de 500x 500 pour obtenir des résultats uniforme. Ensuite, nous la convertissons en niveaux gris. Cette étape est très importante car elle va nous permettre de normaliser les entrées.

1.1.2. Extraction et description des caractéristiques :

Nous allons utilisé deux algorithmes séparés pour extraire les attributs et créer des descripteurs. Après avoir lu les images de la base une par une, et le traiter.

On extrait un vecteur descripteur contenant tous les attributs extraits à toutes les résolutions. Ce vecteur descripteur est stocké dans la base de données.

a. Méthode SURF :

Pour l'extraction des points d'intérêt, nous allons utilisé dans notre système la méthode de caractérisation la plus populaire SURF ; l'idée générale de cette méthode est de transformer une image en vecteurs de caractéristiques, lesquels doivent être dans l'idéal invariants aux transformations géométriques (rotation, mise à l'échelle), et dans une moindre mesure invariants à l'illumination. Il s'agit de détecter des points d'intérêt, qui vont permettre d'identifier un objet. La détection de ces points donne lieu à la mise en place des vecteurs de caractéristiques dont les composantes sont propres au point considéré. [21], Un exemple des points d'intérêt détectés est montré sur la Figure III.2.



Figure III.2 : Exemple points d'intérêt SURF.

Les descripteurs des points clés sont calculés à partir des orientations de gradient dans des sous-régions des voisinages des points.

Les orientations sont calculées à partir de la réponse du voisinage à une série d'ondelettes de Haar, ce qui ne permet pas une description aussi précise de l'angle d'orientation.[21]. Pour cette raison, la taille du descripteur de SURF est 64.

b. Méthode HOG :

Histogramme des dégradés orientés (HOG) est utilisé pour extraire les caractéristiques globale, à cette méthode, le but est de reconnaître l'image comme le groupe d'histogrammes locaux.



Figure III.3 : Les points d'intérêt en vert ; sous-région 4 X 4 en rouge .

La première étape consiste à détecter tous les points d'intérêt de l'image en utilisant le détecteur Harris. Cet opérateur est basé sur la matrice d'auto-corrélation qui décrit la structure locale de l'image.

Calculez ensuite l'histogramme d'orientation du dégradé autour de la région de 16 x 16 pixels de chaque point d'intérêt. Premièrement, la région est divisée en sous-région de 4 X 4. (Figure III.4) Pour chaque sous-région, on calcule l'orientation du gradient de 8 cases $C(N)$, $N = 1$ à 8 qui forme un vecteur caractéristique de dimension 128 ($4 \times 4 \times 8$). Chaque pixel d'une cellule vote pour une orientation entre 0 et 180 dans le cas non signé (Figure III.4), ou entre 0 et 360 dans le cas signé.

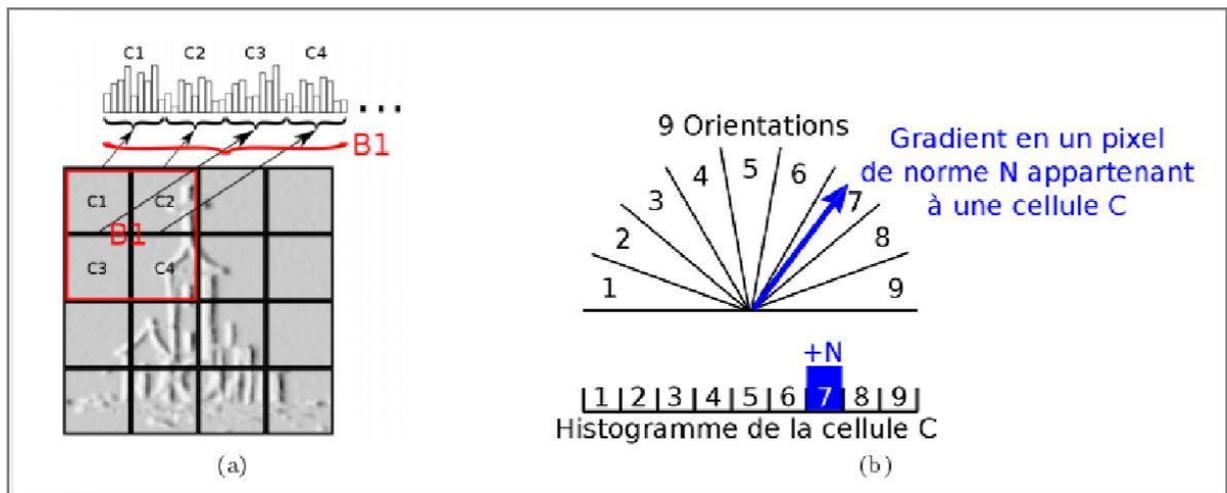


Figure III.4 : Formation du HOG d'une image: (a) découpage en cellules et formation de l'histogramme par bloc, (b) vote d'un pixel selon l'orientation de ses gradients.

Les cases de direction sont définies par l'angle relatif par rapport à la direction de gradient dominante D de la région d'image. Enfin, combiner ensemble l'histogramme d'orientation Gradient de la zone de points d'intérêt pour former un vecteur de caractéristiques de taille 128 dimension.

Le vecteur de HOG pour chaque cellule est normalisé en utilisant des vecteurs dans leurs blocs respectifs. En utilisant la méthode norme L2 pour le bloc de normalisation.[22]

1.1.3. Représentation par des caractéristiques :

Lorsque les attributs sont extraits l'image est représentée, soit directement par la liste de ses attributs locaux (présentés sous forme de vecteurs) soit à partir d'une représentation globale supplémentaire qui s'appuie sur une approche dite par « sac de mots ».

La représentation par sac de mots visuels est une des plus populaires obtenir des représentations globales d'images à partir de caractéristiques visuelles locales. Il se compose de trois étapes principales comme suit :

- 1) Après détection et description des caractéristiques extraites des régions locales
- 2) La deuxième étape consiste à convertir les vecteurs de descriptions dans un dictionnaire de mots (code Words) chaque code de mots résulte d'une méthode de regroupement k -Means et représente ainsi plusieurs régions similaires.

L'algorithme des k -means est généralement utilisé où k représente le nombre de mots du dictionnaire. L'application de l'algorithme k -Means se fait suivant les étapes suivantes:

- Choix de k , le nombre de cluster à créer.

- Choix des centres des clusters de manière aléatoire à partir des objets en entrée.

La procédure adoptée pour le choix des centres des Clusters initiaux est extrêmement importante car elle a un impact direct sur le résultat final du Classification. Il est donc très important de choisir des clusters bien séparés.

L'algorithme K-Means basique se base sur une initialisation aléatoire.

- Parcourir tous les objets afin de les affecter ou les réaffecter au cluster approprié en se basant sur la minimisation de la distance entre l'objet et le centre du cluster.

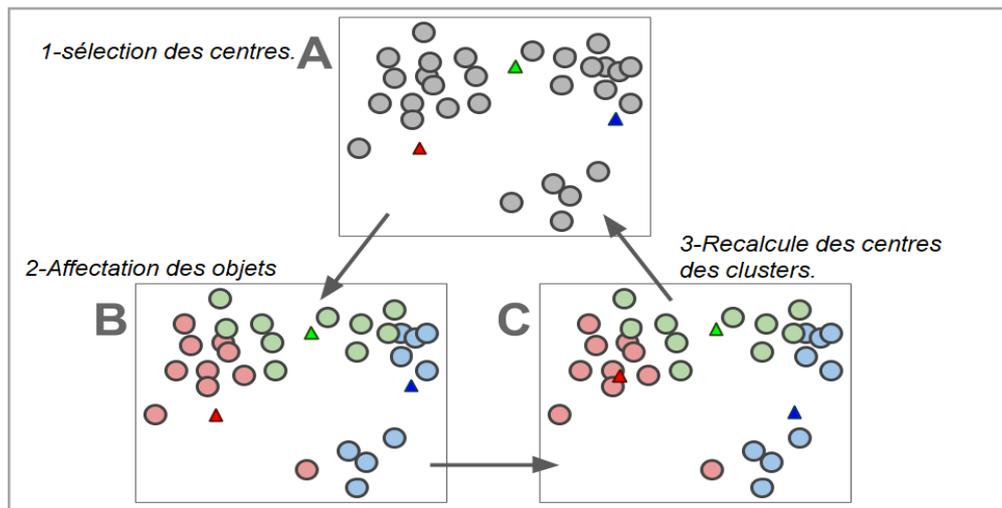


Figure III.5: l'algorithme K-Means.

- Calculer les centres de chaque cluster puisqu'ils peuvent changer après affectation Des objets.

- Refaire les étapes (3) et (4) jusqu'à aucun changement du calcul des centres des Clusters ou une stabilité des objets.

3) La dernière étape consiste à représenter l'image comme un histogramme de mots visuels.

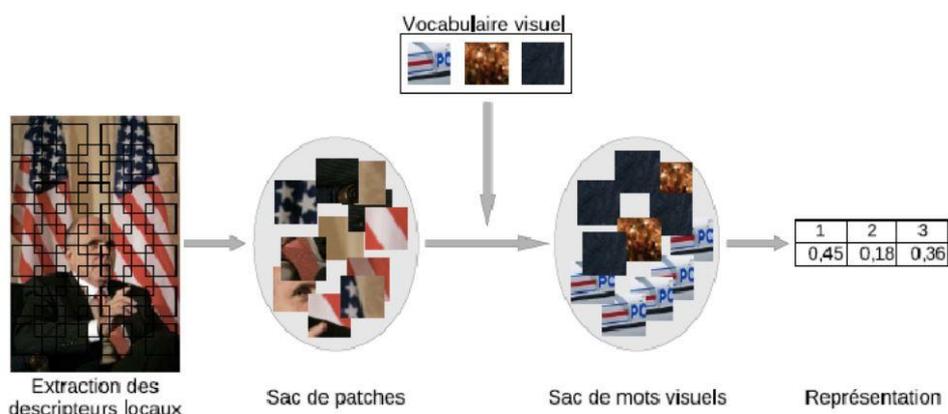


Figure III.6: Représentation en histogramme du vocabulaire visuel

1.1.4. Apprentissage :

L'objectif est d'obtenir une base d'apprentissage à partir de l'application de la méthode SVM. Cette étape prend comme entrée le vecteur de caractéristique de chaque image obtenu de l'étape précédent et l'étiquette de la classe (damage et non damage) qui appartient ce vecteur et parcourt la base d'image pour obtenir la base d'apprentissage.

1.2. Phases de reconnaissance :

Le but de cette étape est de connaître à quelle classe appartient l'image recherchée, donc pour chaque image entrée, notre système passe par les étapes suivantes:

- l'image d'entrée est prétraitée pour l'extraction des caractéristiques.
- Les descripteurs sont calculés à partir de l'image d'entrée.
- Calculer le vecteur Bag-of-Words avec l'algorithme k-Means.
- la classification avec l'algorithme SVM.

La classification est le problème inverse de l'apprentissage. Ça consiste à utiliser l'information issue de l'apprentissage pour attribuer à chaque point ou à chaque région de l'image une classe.

Conclusion :

Nous avons présenté dans ce chapitre les différentes étapes et architecture plus détaillée à une conception convenable d'un système de classification et reconnaissance d'images, nous avons choisi pour l'extraction des caractéristiques deux méthodes séparées HOG et SURF et la méthode SVM pour la reconnaissance des différentes classes.

Dans le prochain chapitre, nous allons illustrer la réalisation de notre système en représentant quelques interfaces et résultats obtenus.

***Chapitre IV: Implémentation
et réalisation***

Introduction :

Ce chapitre présente l'implémentation et la réalisation de notre application. On commence tout d'abord par une présentation du langage de programmation choisi. Ensuite nous présentons les tests réalisés et les discussions du résultat de la méthode d'extraction de descripteurs. On termine ce chapitre par des captures d'écran de l'exécution de l'application.

1. Environnement du développement :

Pour réaliser notre application, on a choisi d'utiliser MATLAB, (matrix laboratory) est un logiciel interactif basé sur le calcul matriciel. Il est utilisé dans les calculs scientifiques et les problèmes d'ingénierie parce qu'il permet de résoudre des problèmes numériques complexes en moins de temps requis par les langages de programmation courants, et ce grâce à une multitude de fonctions intégrées et à plusieurs programmes, outils testés et regroupés selon usage dans des dossiers appelés boîtes à outils ou "Toolbox".

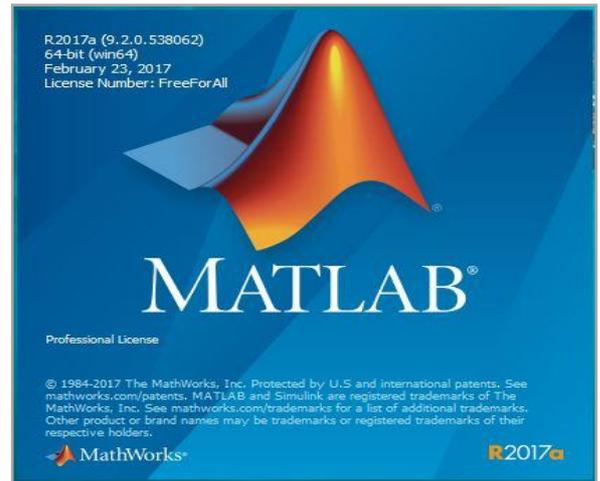


Figure IV.1 : MATLAB

En plus, MATLAB comporte des fonctions prédéfinies embarquées dédiées au traitement d'image.

On a utilisé MATLAB R2017a (9.2.0.538062) une version 64-bit de : 23.02.2017.

2. Base d'images et zone d'étude :

2.1. Base d'images :

Nous avons utilisé deux différents ensembles de données d'images satellitaires pour effectuer des expériences dans ce travail. Il est constitué d'un ensemble de 84 images satellitaires (44 images dommages et 40 images No dommages).

2.2. Zone d'étude :

Deux séismes destructeurs récents ont été considérés pour cette étude ; Haïti et Népal :

- **Séisme en Haïti :**

Un séisme majeur avec une magnitude de (MW 7,0) centré environ 15 milles au sud-ouest de Port-au-Prince, la capitale d'Haïti, a frappé la côte haïtienne le 12 janvier 2010.

Nous avons des images satellites fournies par Google Earth. Dans certains quartiers de la capitale d'Haïti, Port-au-Prince, avant et après le tremblement de terre d'une cathédrale, un

hôtel de luxe à Port-au-Prince et le palais présidentiel d'Haïti, et le stade de football Silvio Cator.

- **Séisme en Népal :**

Un puissant séisme (7,8) a secoué le Népal à proximité de Barpak, entre la capitale Katmandou et la ville Pokhara le dimanche 25 avril 2015.

Nous avons obtenu des images prises par le satellite Pléiades, avant et après le tremblement de terre de Katmandou, Durbar Square, Tour Bhimsen et quartier de Sundhara avant et après le séisme montrent l'impact dévastateur du tremblement de terre.

3. Résultats obtenues :

3.1. La précision :

La précision est le rapport entre le nombre d'images pertinentes retrouvées et le nombre d'images retrouvées, La précision sont compris entre 0 et 1. Cette mesure permet de rendre compte de la pertinence de la réponse du système à la requête de l'utilisateur.

$$Précision = \frac{|Ra|}{|A|}$$

R: l'ensemble d'images pertinentes dans la base d'images utilisée pour évaluer.

|R| : le nombre d'images pertinentes dans la base d'images.

A: l'ensemble des réponses.

|A| : le nombre d'images dans l'ensemble des réponses.

|Ra| : le nombre d'images pertinentes dans l'ensemble des réponses.

3.2. Comparaison des résultats :

Le Tableau. IV.1 montre la variation de la précision obtenue pour chaque K et chaque descripteur.

K_means	BOVW_HOG	BOVW_SURF
30	0.5333	0.600
40	0.400	0.600
50	0.600	0.666
55	0.3333	0.480
60	0.3333	0.466
70	0.3333	0.40

Tableau IV.1 : Variation de la précision de classification

BOVW est basé sur une classification non supervisée (K-means) que nous utilisons pour construire le dictionnaire. Tableau. IV.1 montre la comparaison résulte en faisant varier le nombre de classes (K) pour obtenir le vocabulaire visuel.

Nous pouvons observer que la précision de la classification varie entre 33% et 66%.

Nous notons que pour K=50 dictionnaires, les deux approches fournissent de meilleurs résultats. Mais la précision de SURF est meilleure que celle HOG dans la plupart des cas k.

4. Présentation de l'application :

Les figures ci-dessous illustrent quelques interfaces de notre application de classification des images en figure :

1. Interface utilisateur :

L'interface principale de notre système qui comporte les boutons principaux de l'application (HOG et SURF) :

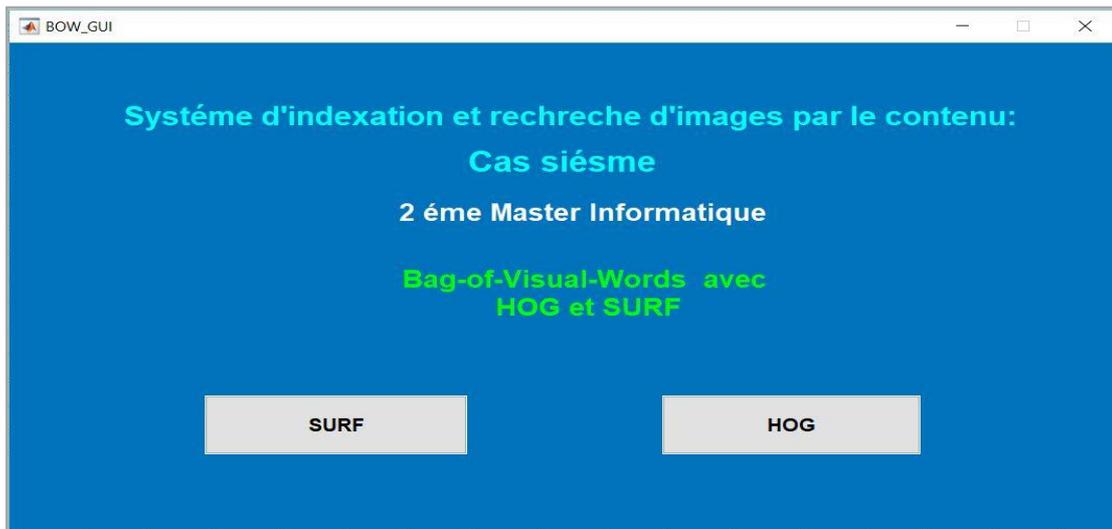


Figure IV.2 :l'interface de démarrage de notre système.

- Si vous clic sur le bouton "SURF" de cette interface permet de lancer notre descripteur (GUI_SURF) et d'afficher une nouvelle fenêtre comme suit:

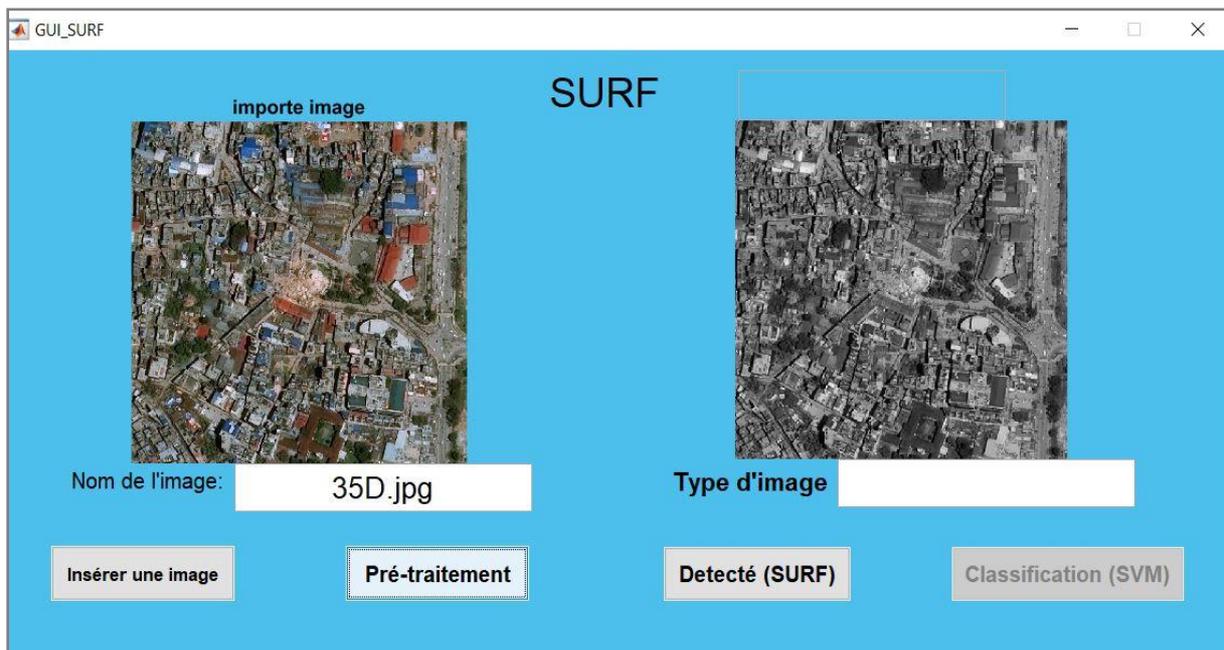


Figure IV.3 : Interface GUI_SURF.

- Bouton '**Insérer une images**' pour insérer une image à partir des fichiers de test
- Bouton '**Pré-traitement**' pour convertir cette image en gris et la redimensionner.
- Bouton '**Déecté(SURF)**' pour appeler la fonction 'SURF' afin d'extraire les points d'intérêt dans l'image et calculer la description .
- Bouton '**Classification(SVM)**' pour classer l'image importée et afficher le résultat dans la boîte de message.

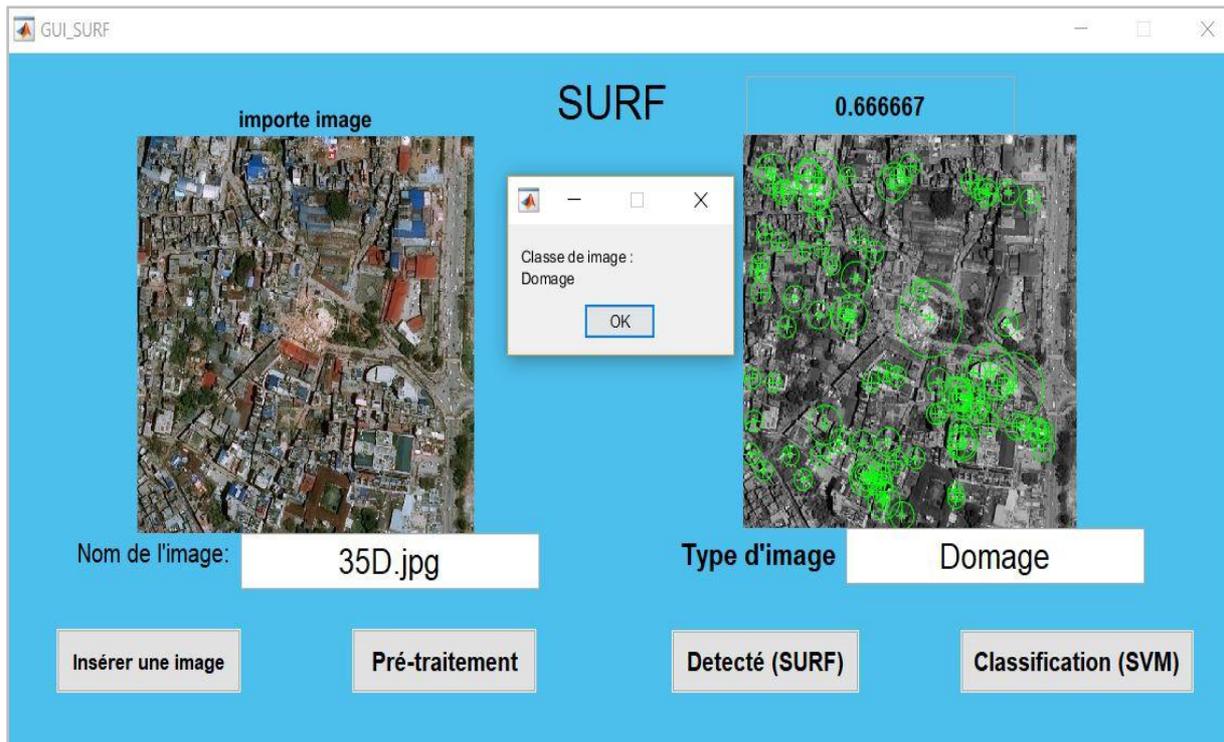


Figure IV.4 : Affichage des résultats de classification(SURF).

➤ **Bouton HOG** : Le clic sur le bouton "HOG" dans l'interface principale permet de lancer notre descripteur HOG(GUI_HOG) et d'afficher une nouvelle fenêtre comme suit:

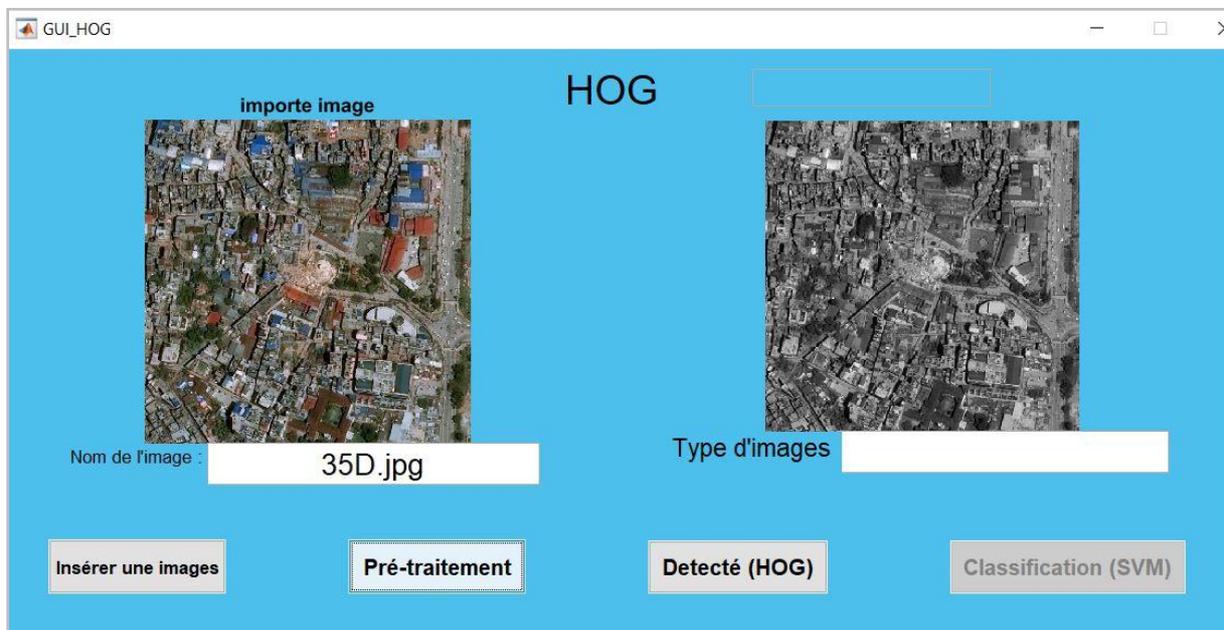


Figure IV.5 : Interface GUI_HOG.

- Bouton '**Insérer une images**' pour insérer une image à partir des fichiers de test
- Bouton '**Pré-traitement**' pour convertir cette image en gris et la redimensionner.

- Bouton '**Défecté(HOG)**' pour appeler extraites caractéristiques du HOG à partir d'une image d'entrée.
- Bouton '**Classification(SVM)**' pour classer l'image importée et afficher le résultat dans la boîte de message.

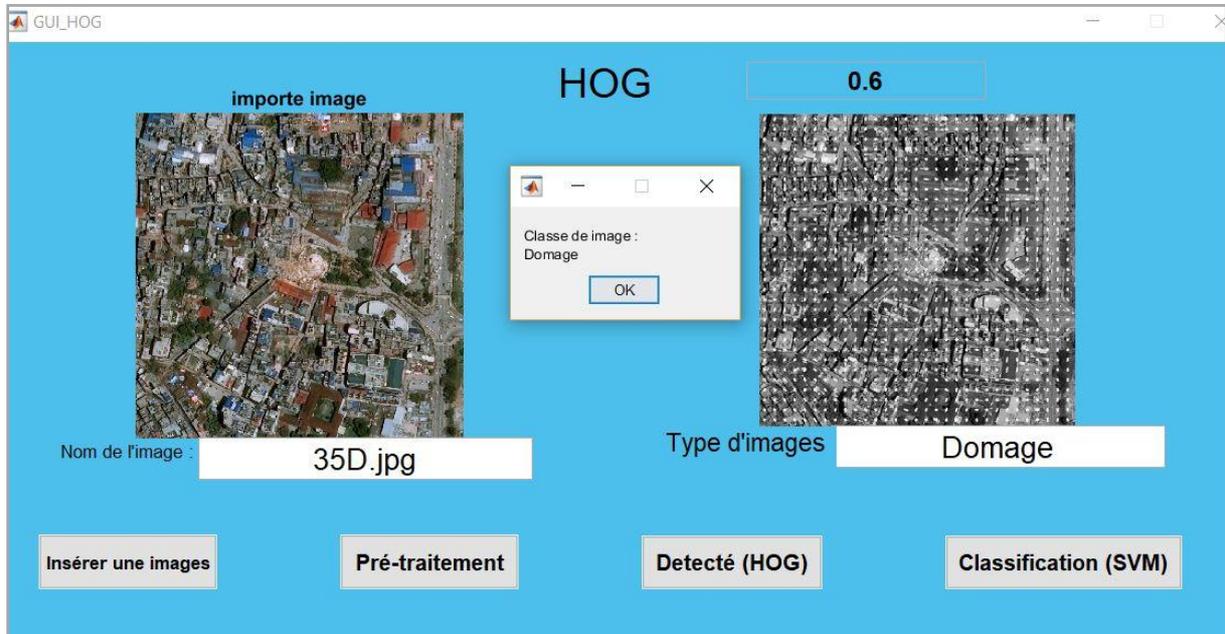


Figure IV.6 : Affichage des résultats de classification(HOG).

Conclusion :

Nous avons présenté dans ce chapitre notre application et travaux pratiques effectués durant cette mémoire, au début on a présenté le software utilisé dans ce mémoire et une mesure d'évaluation avec laquelle on a calculé la précision des résultats, les différents résultats obtenus pendant l'évaluation des différents descripteurs visuels. Notamment sur la représentation par sacs de mots et sur le clustering.

Conclusion générale & perspectives

Dans ce mémoire, nous avons abordé le problème de classification des images satellitaires. Plus précisément, nous nous sommes focalisés sur la classification basée sur le contenu visuel. Nous proposons un système de classification d'images satellitaires basé sur l'extraction de caractéristiques locales et globales pour reconnaissance et détection automatiques des bâtiments endommagés.

Le premier descripteur utilisé est SURF (Speeded Up Robust Features) ; le recours à ce descripteur se justifie par les résultats intéressants obtenus dans des travaux de reconnaissances d'objets il représente une solution plus performante en terme de rapidité que le détecteur SIFT.

Le deuxième descripteur utilisé est HOG (Histogram of Gradient Orientation). L'avantage de ces descripteurs est la simplicité de leur mise en œuvre, ainsi que le nombre réduit d'observations que l'on obtient.

En comparant les résultats obtenus de ces deux descripteurs, nous avons constaté que le descripteur SURF à donner de bons résultats en utilisant une représentation compacte d'images (représentation par les sacs de mots visuels).

Après avoir décrit les images à l'aide d'ensembles de descripteurs, nous les avons représentés avec des dimensions plus réduite par les sacs de mots visuels. Ensuite nous avons employé l'approche clustering K_means afin de limiter le nombre d'images à vérifier.

On a choisi aussi la méthode de SVM, comme méthode de classification, ce choix est justifié par leur principal avantage pour tous les méthodes supervisées, qui est l'aspect d'apprentissage.

Perspectives et travaux futurs :

Ce travail est amené à être prolongé dans plusieurs directions :

- L'utilisation d'une grande base d'image dans l'apprentissage pour augmenter les performances de la classification.
- L'ajout ou la combinaison d'autres méthodes d'extraction de caractéristiques pour mieux présenter les propriétés, afin d'améliorer le taux de classification.

- Une étude comparative d'autres méthodes d'apprentissage statistique peut amener à sélectionner le meilleur modèle approprié à la détection.

En fin, on peut dire que ce travail reste ouvert pour des travaux de comparaison et/ou d'hybridation avec d'autres méthodes de classification et d'autres types de descripteurs.

Bibliographies:

- [01]: Claude. K: *«la télédétection aérospatiale: une introduction »*, Cours Ecole Normale Supérieure Paris.
- [02]: Youssef EL MERABET: *«Segmentation d'images couleur par combinaison LPE-régions/LPE-contours et fusion de régions - Application à la segmentation de toitures à partir D'orthophotoplans »* Thèse de Doctorat, université de technologie Belfort – Montbéliard.
- [03]: Antoine .L.:*«Contribution de la texture pour l'analyse d'images à très haute résolution spatiale. Application à la détection de changement en milieu périurbain »*, Thèse pour obtenir docteur de l'université européenne de Bretagne, 2011.
- [04]: Chesnel.A : *«Quantification de dégâts sur le bâti lies aux Catastrophes majeures par images satellite Multimodales très haute résolution »*, thèse pour obtenir docteur, 2008.
- [05]: Simon. E: *«Mise en forme d'une base de données sur le séisme en Haïti et exploitation de données Lidar post-événement pour caractériser le bâti et évaluer les dommages »* Mémoire de soutenance de Diplôme d'Ingénieur, 23 Septembre 2011.
- [06] : Soumia. B: *« Indexation et Recherche d'Images par Contenu dans les Grands Corpus d'Images »* Thèse Docteur de l'Université d'Oran, Année 2017.
- [07] : M.A.Bourenane : *«Un outil pour l'indexation des vidéos personnelles par le contenu »*. Thèse de Doctorat, Université de Québec à trois- rivières. 2009.
- [08]:Gwénélel.Q : *«Indexation et fusion multimodale pour la recherche d'information par le contenu. Application aux bases de données d'images médicales »*.Thèse de Doctorat, Université européenne Bretagne.Septembre 2008.
- [09]: K.Houari:*«recherche d'images par le contenu »*,Thèse De Doctorat,2010.
- [10]:D.Mercier,R.Séguier : *« Utilisation des STANN en audio : illustrationen reconnaissance de chiffre »*, Journée Valgo 2001.
- [11]: DU Peijun ,CHEN Yunhao,TANG Hong et FANG Tao :*«Study on Content-Based Remote Sensing Image Retrieval»* ,2015.

- [12]: N.Dalal et B.Triggs : « *Histograms of Oriented Gradients for Human Detection* ». INRIA Rhone-Alps, Computer Vision and Pattern Recognition Conference 2005.
- [13]: BEDOUHENE.S:«*Recherche d'images par le contenu*»,memoire de magister universite tizi-ouzou.
- [14]:S.PIERARD:«*Développement d'une méthode de représentation d'images sur base d'un vocabulaire visuel*».
- [15]:J.LANDRÉ: «*analyse multi résolution pour la recherche et l'indexation D'images par le contenu dans les bases de données images Application à la base d'images paléontologique*», Thèse De Doctorat.
- [16] :W.BOUACHIR :«*Recherche d'images et classification par mots visuels et descripteurs flous*».Université de Moncton, avril 2010.
- [17]: R.Arandjelovic et A.Zisserman : «*All about VLAD*».
- [18]: G.Amato et F.Falchi et L.Vadicamo : «*Aggregating binary local descriptors for image retrieval* »,2017.
- [19]: H.Jegou et M.Douze et C.Schmid et P.Pérez :«*Aggregating local descriptors into a compact image representation*».
- [20]: M.Issolah : « *Identification d'images pour la botanique* ».
- [21]: T.Baerecke, M.Bendris : «*Traitement des modalités « image » et « vidéo »* ».
- [22]: Manuel Grand-brochier : «*Descripteurs 2D et 2D+t de points d'interet pour des appariements robustes*». Thèse Docteur ,2011.