

UNIVERSITE SAAD DAHLEB DE BLIDA

Faculté des sciences

Département d'informatique



MEMOIRE DE MASTER

En Informatique

Option : Ingénierie Des Logiciels

THÈME :

**Classification d'images pour la
gestion des déchets solides
municipaux**

Réalisé par
MOUSSERATI Fatma Zohra
OULD SETTI Hiba

Encadré par
Dr. BACHA Siham
Dr. AOUACHE Mustapha

18 Juillet 2019

Remerciements

Nous remercions ALLAH de nous avoir donné la santé et le courage afin de pouvoir réussir ce travail. Ce travail est l'aboutissement d'un long cheminement au cours duquel nous avons bénéficié d'encadrement, des encouragements et du soutien de plusieurs personnes, à qui nous tenons à dire profondément et sincèrement merci.

Nous exprimons notre grande gratitude à notre promotrice Mme. BACHA Siham, d'avoir accepté de nous encadrer durant notre travail et pour ses précieux conseils et ses orientations.

Nous remercions très chaleureusement Mr. AOUACHE Mustapha pour son soutien, ses précieux conseils tout au long de notre stage au niveau du CDTA. Nous avons eu le privilège de travailler parmi votre équipe et d'apprécier vos qualités et vos valeurs, votre sérieux et votre compétence.

Nous désirons exprimer notre reconnaissance envers Mr RIALI Ishak qui nous a apporté son support moral et intellectuel tout au long de notre démarche.

Nous remercions également tous les membres du jury pour nous avoir honorés par leur présence et pour avoir accepté d'évaluer notre travail.

Enfin nous tenons à remercier l'ensemble du corps enseignant, en particulier Mme Rezzoug et Melle Ouahrani que Dieu la bénisse. Nous gardons un très bon souvenir de vous, nous apprécions toujours vos qualités humaines, votre savoir-faire et vos compétences scientifiques.

Résumé

De nos jours, la gestion des déchets urbains représente un des problèmes majeurs pour les municipalités algériennes. La multiplication des décharges sauvages le long des rues, et l'entassement de grandes masses d'ordures ménagères à l'air libre contribuent à détériorer notre milieu. La mise en œuvre d'un système intelligent pour la gestion des déchets permet de faciliter, aux services publics et aux entreprises concernées, le nettoyage, la collecte et l'élimination des déchets.

Ainsi devant la vivacité du problème et l'importance de son enjeu social et environnemental, nous avons adopté une méthodologie de recherche qui consiste à concevoir et implémenter deux grands systèmes de classification automatique d'images. Le but de ces systèmes est de classer les images de poubelles en fonction de leurs niveaux de remplissage (vide, à moitié remplie, ...). Les systèmes utilisent un algorithme de classification supervisé pour analyser les caractéristiques visuelles extraites à partir des images de poubelles et identifier leurs taux de remplissage.

Mots clés :

Classification d'images, Apprentissage automatique, gestion des déchets solides, extraction de caractéristiques.

Abstract

Today, solid waste management is one of the major problems for Algerian municipalities. The accumulation of large masses of garbage in the open air contribute to damage our environment. The implementation of an intelligent system for waste management facilitates the cleaning, collection and disposal of waste by the public services and the companies concerned. So in front of this problem and the importance of its social and environmental stake, we have adopted a research methodology that consists of designing and implementing two major automatic image classification systems. The purpose of these systems is to classify garbage bin images according to their fill levels (Low, Full, Flow, Medium, OverFlow). The systems use a supervised classification algorithm to analyze the visual characteristics extracted from the bin images and to identify their fill rates.

Keywords :

Image classification, Machine learning, Solid waste management, features extraction.

Table des matières

Table des figures	10
Liste des tableaux	12
Liste des abréviations	15
Introduction Générale	17
Contexte de travail	17
Problématique	17
Objectifs du travail	18
Organisation du mémoire	18
1 La gestion des déchets solides	19
1.1 Introduction	19
1.2 Définition du déchet :	19
1.3 Classification des déchets :	19
1.3.1 Distinction en fonction de l'activité à l'origine du déchet	20
1.3.2 Distinction en fonction de la nature du déchet :	20
1.3.3 Distinction en fonction du mode (filière) de traitement des déchets :	21
1.4 Système de gestion des déchets :	21
1.5 Stratégies de la gestion des déchets :	23
1.5.1 Les objectifs de la gestion des déchets [14]	23
1.5.2 Les choix stratégiques	24
1.6 Impacts des déchets solides sur la santé et l'environnement[22]	24
1.6.1 Impacts sur la santé humaine	24
1.6.2 Impacts sur l'environnement :	25
1.7 La gestion des déchets solides en Algérie	25
1.7.1 Principes de la gestion de déchets en Algérie [30]	25
1.8 Revue sur quelques travaux connexes	27

1.9	Conclusion	28
2	Techniques de classification d'images	29
2.1	Système d'indexation et de recherche d'images	29
2.1.1	Principe général de l'indexation de la recherche d'information	30
2.1.2	Concepts fondamentaux de la recherche d'image	31
2.1.3	Architecture des systèmes d'indexation par contenu	32
2.1.4	Principales techniques d'indexation par contenu	33
2.1.5	Les mesures de similarité	35
2.2	Techniques d'apprentissage automatique	35
2.2.1	Apprentissage supervisé	35
2.2.2	Apprentissage non-supervisé	36
2.2.3	Apprentissage profond	36
2.2.4	Apprentissage par transfert	44
2.3	La validation croisée	48
2.4	Conclusion	49
3	Système de classification d'images pour la gestion des déchets	51
3.1	Introduction	51
3.2	Contexte et motivation	51
3.3	Solution à base de système de recherche classique	52
3.3.1	Phase 1 : Module d'indexation d'images	53
3.3.2	Phase 2 : L'indexation de la requête	56
3.3.3	Phase 3 : Module de classification	57
3.4	Solution à base de système de recherche adaptative	60
3.4.1	Méthodologie et conception du système de classification basé sur le fine-tuning	60
3.4.2	Méthodologie et conception du système à base d'apprentissage profond pour l'extraction des caractéristiques	64
3.5	Conclusion	66
4	Résultats expérimentaux et évaluation	67
4.1	Introduction	67
4.2	Démarche expérimentale	67
4.2.1	Ressources matérielles et logicielles utilisées lors du développement	67
4.2.2	Outils développés	69
4.3	Présentation de la base de données et métriques d'évaluation	73
4.3.1	Présentation de la base de données	73
4.3.2	Métriques d'évaluation	74
4.4	Résultats et discussion	76

4.4.1	Résultats du système de recherche classique	76
4.4.2	Résultats du système de recherche adaptative	82
4.4.3	Comparaison avec l'état de l'art	90
4.5	Conclusion	91
	Conclusion et perspectives	93
	Bibliographie	95

Table des figures

2.1	Processus de RI	30
2.2	Architecture d'un système d'indexation et de recherche d'images	33
2.3	L'architecture générale d'un Perceptron [29].	37
2.4	Schéma démonstratif de la fonction d'activation [58].	37
2.5	Graphe de la fonction Heaviside [29].	38
2.6	L'architecture générale d'un MLP [18].	39
2.7	Les fonctions d'activation utilisés par les réseaux de neurones [37].	39
2.8	Differents taux d'apprentissage [51].	41
2.9	L'architecture globale d'un réseau de neurones convolutifs [19].	42
2.10	un simple réseau de neurones avec 3 couches vs un CNN avec 3 couches en 3 dimensions [1].	42
2.11	Application de plusieurs filtres sur une image d'entrée [19].	43
2.12	l'apprentissage traditionnel et l'apprentissage par transfert [53].	44
2.13	Les trois mesures qui peuvent améliorer l'apprentissage par le transfert [59].	45
2.14	Extraire les caractéristiques avec un CNN [?].	45
2.15	l'architecture de VGG-16 [54].	46
2.16	l'architecture de ResNet50 [34].	47
2.17	l'architecture de MobileNet [36].	47
3.1	L'architecture de l'approche proposée à base de système de recherche classique	52
3.2	Exemple illustre l'effet d'ajustement du contraste	53
3.3	Processus d'extraction des caractéristiques à partir des images	54
3.4	Quantification d'une image	55
3.5	Exemple du GLCM appliqué sur une image de la base d'images du système	56
3.6	Exemple de quelques vecteurs de caractéristiques	56
3.7	Schéma détaillé de la phase de classification	58
3.8	L'architecture de l'approche proposée basée sur la technique de fine-tunnig.	61
3.9	Exemple d'une image avant et après le pré-traitement	61

3.10	Processus de la classification d'une requête.	63
3.11	L'architecture de l'approche proposée à base d'apprentissage profond pour l'extraction des caractéristiques.	64
3.12	Processus de la classification d'une requête.	66
4.1	Références de l'ordinateur personnel utilisé	67
4.2	Interface d'accueil du système classique	70
4.3	Classification de l'image requête	70
4.4	Interface des images similaires à l'image requête	71
4.5	Interface d'accueil du système adaptatif	72
4.6	Interface d'affichage de la classe de l'image requête	72
4.7	Interface des images similaires à l'image requête	73
4.8	Exemple d'images de la base de données des différentes classes.	74
4.9	Matrice de confusion [3].	75
4.10	Matrice de confusion du 3 plus proches voisins avec la distance Euclidien	77
4.11	Matrice de confusion du 3 plus proches voisins avec la distance Euclidien avec la validation croisée (14 partitions)	78
4.12	Matrice de confusion du 3 plus proches voisins avec la distance Corrélation sans la validation croisée	81
4.13	Matrice de confusion du 3 plus proches voisins avec la distance Corrélation et la validation croisée (14 partitions)	82
4.14	Matrice de confusion du modèle ResNet-50 sans validation croisée.	84
4.15	Matrice de confusion du modèle ResNet-50 avec validation croisée.	84
4.16	Matrice de confusion du modèle MobileNet sans validation croisée.	85
4.17	Matrice de confusion du modèle MobileNet avec validation croisée.	85
4.18	Matrice de confusion du modèle VGG-16 sans validation croisée.	86
4.19	Matrice de confusion du modèle SVM sans validation croisée.	87
4.20	Matrice de confusion du modèle SVM avec validation croisée pour $k=2$	88
4.21	Matrice de confusion du modèle SVM avec validation croisée pour $K=3$	88
4.22	Matrice de confusion du modèle SVM avec validation croisée pour $K=5$	89
4.23	Matrice de confusion du modèle SVM avec validation croisée pour $K=7$	90

Liste des tableaux

3.1	Application numérique des descripteurs statistiques	56
3.2	Nombre d'occurrence des classes	58
4.1	Résultats du K-ppv avec 8 niveaux de gris et la distance Euclidien	76
4.2	Résultats du K-ppv avec 16 niveaux de gris et la distance Euclidien	76
4.3	Résultats du K-ppv avec 8 niveaux de gris et la distance Euclidien (avec 7 partitions)	77
4.4	Résultats du K-ppv avec 16 niveaux de gris et la distance Euclidien (avec 7 partitions)	77
4.5	Résultats du K-ppv avec 8 niveaux de gris et la distance Euclidien (avec 14 partitions)	78
4.6	Résultats du K-ppv avec 16 niveaux de gris et la distance Euclidien	78
4.7	Résultats du K-ppv avec 8 niveaux de gris et la distance Cosine	79
4.8	Résultats du K-ppv avec 8 niveaux de gris et la distance Cosine	79
4.9	Résultats du K-ppv avec la quantification n=8 et la distance Cosine (avec 7 partitions)	79
4.10	Résultats du K-ppv avec 16 niveaux de gris et la distance Cosine (avec 7 partitions)	79
4.11	Résultats du K-ppv avec 8 niveaux de gris et la distance Cosine (avec 14 partitions)	79
4.12	Résultats du K-ppv avec 16 niveaux de gris et la distance Cosine (avec 14 partitions)	80
4.13	Résultats du K-ppv avec 8 niveaux de gris et la distance Corrélacion	80
4.14	Résultats du K-ppv avec 8 niveaux de gris et la distance Corrélacion	80
4.15	Résultats du K-ppv avec 8 niveaux de gris et la distance Corrélacion(avec 7 partitions)	81
4.16	Résultats du K-ppv avec 16 niveaux de gris et la distance Corrélacion(avec 7 partitions)	81
4.17	Résultats du K-ppv avec 8 niveaux de gris et la distance Corrélacion (avec 14 partitions)	82
4.18	Résultats du K-ppv avec 16 niveaux de gris et la distance Corrélacion(avec 14 partitions)	82
4.19	Résultats de fine-tuning.	83
4.20	Résultats de fine-tuning avec la validation croisée.	83
4.21	Résultats du classificateur SVM avec VGG-16, MobileNet, ResNet-50 sans validation croisée.	86
4.22	Résultats du classificateur SVM avec VGG-16, MobileNet, ResNet-50 avec la validation croisée pour K=2.	87
4.23	Résultats du classificateur SVM avec VGG-16, MobileNet, ResNet-50 avec la validation croisée pour K=3.	87

4.24 Résultats du classificateur SVM avec VGG-16, MobileNet, ResNet-50 avec la validation croisée pour K=5.	88
4.25 Résultats du classificateur SVM avec VGG-16, MobileNet, ResNet-50 avec la validation croisée pour K=7.	89
4.26 Évaluation par rapport aux travaux connexes	90

Liste des abréviations

- ADEM** Agence de l'environnement et de maîtrise d'énergie. 26
- AND** Agence Nationale des Déchets. 26
- ANN** Artificiel Neural Network. 38
- APC** Assemblée Populaire Communale. 27
- API** Interface de Programmation Applicative. 69
- CBIR** Content-Based Image Retrieval. 29, 51
- CET** Centre d'Enfouissement Technique. 26
- CNFE** Centre National de Formation en Environnement. 27
- CNN** Convolutional Neural Network. 10, 41, 42, 45, 62, 65, 90
- FN** False Negative. 74
- FP** False Positive. 74
- GLCM** Gray Level Co-occurrence Matrix. 10, 28, 33, 34, 52, 53, 55, 56, 90
- K-ppv** K-plus proches voisins. 12, 28, 52, 57, 76–82, 90
- MATET** Ministère de l'Aménagement du Territoire de l'Environnement et du Tourisme. 26
- MLP** Multi Layer Perceptron. 38, 90
- Pixel** Picture Element. 31, 33, 34, 73
- PROGDEM** Programme national de gestion des déchets ménagers. 25, 26
- RGB** Red Green and Blue. 31, 53, 61, 62, 64
- RI** Recherche d'information. 10, 30, 31
- SQL** Structured Query Language. 32
- SVM** Support Vectors Machine. 65, 86, 90
- TN** True Negative. 74
- TP** True Positive. 74

Introduction Générale

Contexte de travail

Dans le monde entier, les taux de production de déchets augmentent [61]. En 2016, les villes du monde ont généré 2,01 milliard de tonnes de déchets solides, ce qui correspond à une empreinte au sol de 0,74 kg par personne et par jour [61]. Avec la croissance démographique rapide et l'urbanisation, la production annuelle de déchets devrait augmenter de 70% par rapport aux niveaux de 2016 pour atteindre 3,40 milliards de tonnes en 2050 [61]. Comparés à ceux des pays développés, les résidents des pays en développement, en particulier les citoyens pauvres, sont plus gravement touchés par les déchets gérés de manière non durable. Dans les pays à faible revenu, plus de 90% des déchets sont souvent jetés dans des décharges non réglementées ou brûlés ouvertement. Ces pratiques entraînent des conséquences graves pour la santé, la sécurité et l'environnement. Les déchets mal gérés servent de terrain fertile pour les vecteurs de maladies, contribuent au changement climatique mondial par la génération de méthane et peuvent même promouvoir la violence urbaine.

En Algérie, selon une étude publiée en 2015 [25], des experts Coréens, ayant observé les wilayas de Blida et de Bordj Bou Arreridj, révèlent un faible taux de collecte en raison de l'insuffisance des bacs de récupération des déchets, de l'irrégularité des fréquences de collecte, ainsi que du manque ou de la vétusté des moyens de collecte. L'Algérie devrait faire de la gestion des déchets, l'un des grands enjeux dans le moyen terme notamment avec l'augmentation de sa population qui devrait atteindre les 51 millions d'habitants à l'horizon 2035 [55].

La gestion appropriée des déchets est essentielle pour la construction de villes durables et viables, mais cela reste un défi pour de nombreux pays et villes en développement.

Problématique

Devant la vivacité du problème, plusieurs mesures ont été prises pour améliorer le système de gestion des déchets solides, mais plusieurs problèmes sont couramment rencontrés dans ces systèmes. Une collecte insuffisante de déchets, entraîne le stockage d'une quantité énorme de déchets à l'endroit où ils sont générés ou de rester dans les points de collecte pendant des durées

prolongées, peut entraîner de nombreux risques pour l'environnement et la santé, tels que la poussière, les odeurs et les fumées de combustion [41]. Cependant, le simple fait de collecter les déchets à divers endroits ne résout pas le problème. Il est également nécessaire de disposer d'une estimation des déchets générés dans la corbeille afin de fournir à la direction des données précises pour la planification future. Un système efficace de gestion des déchets solides est nécessaire pour améliorer la santé et la sécurité des personnes [10]. En général, un système efficace de gestion des déchets solides devrait inclure la collecte et le transport des déchets et / ou la détection au niveau de la corbeille [10].

Dans ce travail, nous avons tenté de proposer un système efficace de détection de niveaux des déchets solides pour la gestion des déchets solides municipaux. Ce dernier, est basé sur la classification d'images pour l'extraction des caractéristiques visuelles.

Objectifs du travail

Le travail réalisé dans le cadre de ce projet consiste à :

- Faire un état de l'art sur les systèmes de gestion des déchets solides basés sur la classification d'images de poubelles.
- Étudier plusieurs descripteurs de bas niveau pour représenter l'image.
- Explorer et comparer plusieurs algorithmes de classification.
- Concevoir et mettre en œuvre le système qui permettra de reconnaître le taux de remplissage d'une poubelle.
- Évaluer les performances du système proposé.

Organisation du mémoire

Pour mener à bien notre recherche, nous avons organisé notre travail comme suite :

- Chapitre 1 : Sera consacré aux techniques de classification d'images.
- Chapitre 2 : Présentera état de l'art sur les systèmes de gestion des déchets solides basés sur la classification d'images de poubelles.
- Chapitre 3 : Portera sur la conception du système de classification d'images et les approches proposées.
- Chapitre 4 : Dédié à l'évaluation des performances, ainsi que les différents résultats obtenus.

Ce mémoire sera finalisé par une conclusion générale reprenant les principaux points abordés, et en exposant certaines perspectives d'approfondissement, d'élargissement et d'amélioration de notre travail.

Chapitre 1

La gestion des déchets solides

1.1 Introduction

Les déchets sont l'un des meilleurs indicateurs de la vitalité économique et du mode de vie d'une société. La croissance de la production et le développement, toujours plus rapide, de biens de consommation ont changé en profondeur la structure socio-économique des pays, une société de gaspillage est née. La protection de l'environnement devient de plus en plus une préoccupation collective. La question des déchets est quotidienne et touche chaque être humain. En qualité de consommateur, producteur, éboueur et trieur de déchets recyclables, citoyen ou autorité, chacun peut et doit être acteur d'une meilleure gestion des déchets [14].

1.2 Définition du déchet :

Le déchet, substance d'apparence peu noble, véhicule des dangers d'infection non négligeables. Un déchet est un objet en fin de vie ou une substance issue d'un processus, jugés devenus inutiles ou dangereux ou encombrants, et dont nous voulons se débarrasser. Ou plus simplement, un bien devenu inutile dont le propriétaire cherche à se défaire. La Convention de Bâle sur le contrôle des mouvements transfrontaliers de déchets dangereux définit les déchets comme les « substances ou objets qu'on élimine, qu'on a l'intention d'éliminer ou qu'on est tenu d'éliminer en vertu des dispositions du droit national » [14].

1.3 Classification des déchets :

Les déchets peuvent être solides, liquides ou gazeux. Ils peuvent être dangereux ou non dangereux. Ils peuvent aussi être classés selon leur provenance ou leur composition et leurs propriétés physiques et chimiques. Nous pouvons classer les déchets selon plusieurs critères : en fonction de l'activité à l'origine du déchet, en fonction de la nature du déchet, en fonction du mode (filière) d'élimination du déchet [30].

1.3.1 Distinction en fonction de l'activité à l'origine du déchet

- (a) Les déchets ménagers et assimilés : Les déchets ménagers et assimilés sont les déchets produits par les ménages, les commerçants, les artisans, et même les entreprises et industries quand ils ne présentent pas de caractère dangereux ou polluant : papiers, cartons, bois, verre, textiles, emballages. Ces déchets sont collectés par la commune si, eu égard à leurs caractéristiques et aux quantités produites, ils peuvent être éliminés sans sujétions techniques particulières et sans risques pour les personnes ou l'environnement.
- (b) Les déchets industriels : Les déchets industriels non dangereux ou "banals". Ce sont les déchets des entreprises non dangereux qui sont aussi appelés "déchets assimilés aux déchets ménagers". Les déchets industriels dangereux ou "spéciaux". Ce sont les déchets des entreprises qui, en raison de leurs propriétés dangereuses, ils ne peuvent pas être déposés dans des installations de stockage recevant d'autres catégories de déchets.
- (c) Les déchets agricoles : Sont les déchets issus de l'activité agricole. Il s'agit essentiellement de déchets organiques comme les déjections des animaux, les déchets de culture, etc.
- (d) Les déchets d'activité de soins de santé : se sont les déchets issus des hôpitaux et les autres établissement de soins, les laboratoires et les centres de recherches, les morgues et les centres d'autopsie, les banques de sang et les services de collecte de sang.

1.3.2 Distinction en fonction de la nature du déchet :

- (a) Les déchets dangereux : Les déchets sont considérés comme dangereux s'ils présentent une ou plusieurs des propriétés suivantes : explosif, comburant, inflammable, irritant, nocif, toxique, cancérigène, corrosif, infectieux, toxique pour la reproduction, mutagène, écotoxique.
- (b) Les déchets toxiques en quantités dispersées : Sont des déchets dangereux produits en petites quantités par les ménages, les commerçants ou les PME (garages, coiffeurs, laboratoires photo, imprimeries, laboratoires de recherche...). Il peut s'agir de déchets banals souillés (chiffons, cartons,...), piles, résidus de peinture, etc.
- (c) Les déchets non dangereux : Les déchets non dangereux sont les déchets qui ne présentent aucune des caractéristiques relatives à la "dangerosité" mentionnées ci-dessus (toxique, explosif, corrosif, ...). Ce sont les déchets "banals" des entreprises, commerçants et artisans (papiers, cartons, bois, textiles...) et les déchets ménagers.
- (d) Les déchets inertes : Sont des solides minéraux qui ne subissent aucune transformation physique, chimique ou biologique importante : pavés, gravats, carrelage. Ils proviennent des chantiers du bâtiment et des travaux publics, mais aussi des mines et des carrières.
- (e) Les déchets ultimes : Est ultime un déchet, résultant ou non du traitement d'un déchet, qui n'est plus susceptible d'être traité dans les conditions techniques et économiques du

moment, notamment par extraction de la part valorisable ou par réduction de son caractère polluant ou dangereux.

1.3.3 Distinction en fonction du mode (filère) de traitement des déchets :

- (a) Les déchets biodégradables ou composables : comme les résidus verts, boues d'épuration des eaux, restes alimentaires.... Ces déchets sont au moins pour partie détruits naturellement, plus ou moins rapidement, en général par les bactéries, champignons et autres micro-organismes et/ou par des réactions chimiques laissant des produits de dégradation identiques ou proches de ceux qu'on peut trouver dans la nature, parfois néanmoins contaminés par certains résidus.
- (b) Les déchets recyclables : comme le verre, métaux, matières plastiques. Ces déchets peuvent être réutilisés tels quels dans d'autres domaines ou recyclés.
- (c) Les déchets ultimes : les déchets qui ne sont plus susceptibles d'être traités dans les conditions techniques et économiques du moment. Eux seuls devraient être mis en décharge.
- (d) Les déchets spéciaux et déchets industriels spéciaux : dont font partie les déchets toxiques, les déchets radioactifs et déchets nucléaires qui doivent faire l'objet d'un traitement tout à fait particulier en raison de leur nocivité particulière liée à la radioactivité.

1.4 Système de gestion des déchets :

La gestion des déchets est l'ensemble des opérations et moyens mis en œuvre pour limiter, valoriser ou éliminer les déchets. Opérations de prévention, de réduction à la source, opérations de tri, de pré collecte, de collecte et de transport, opérations de réutilisation, de recyclage et de valorisation, opérations de traitement, jusqu'au stockage des résidus, y compris la surveillance de ces opérations et celle des sites d'élimination de ces déchets.

- (a) La prévention : La prévention des déchets consiste à prévenir ou réduire la production et la nocivité des déchets. La prévention inclut toutes les étapes du cycle de vie d'un produit avant qu'il soit considéré comme un déchet. La prévention quantitative concerne la diminution de la masse et du volume des déchets, la prévention qualitative concerne la nocivité des déchets. Trois stades sont à considérer :
 - La réduction à la source ou « Eco-conception » : Cette étape concerne les industriels. C'est l'étape de prévention la plus en amont dans le cycle de vie du produit. Les industriels doivent réduire dès la conception, la fabrication et jusqu'à la distribution la quantité des matières utilisées. Une prévention qualitative doit également permettre de réduire la toxicité des futurs déchets par un choix judicieux des matières employées.

Il faut donc favoriser la conception d'objets durables, réparables, recyclables et réutilisables.

- La responsabilité élargie du consommateur ou « Eco-consommation » : A ce niveau, le consommateur fait des choix et agit au moment de l'achat (choix du moindre emballage), de l'utilisation et de la gestion des biens qu'il consomme.
- Une seconde vie pour nos vieux objets : Il est parfois possible de réutiliser ou de réemployer certains objets, évitant qu'ils ne deviennent des déchets.

(b) Le réemploi, la réutilisation, la régénération et le recyclage : Le réemploi : Nous employons un déchet pour un usage analogue à son premier emploi. C'est par exemple le réemploi de bouteilles consignées. C'est, en quelque sorte, prolonger la durée de vie du produit avant qu'il ne devienne un déchet.

- La réutilisation : Nous employons un déchet pour un usage différent de son premier emploi.
- La régénération : Par un procédé physique ou chimique, nous redonnons au déchet des caractéristiques qui lui permettent de remplacer une matière première neuve.
- Le recyclage : Nous obtenons à partir des déchets, des matières premières entrant dans la fabrication de nouveaux produits.

(c) Le tri : Opération de séparation des différents flux de déchets par les producteurs. Le tri sélectif consiste à récupérer séparément les déchets selon leur nature. Nous parlons aussi de tri des déchets ou de collecte sélective. Il est possible de faire le tri à la source de création de déchet ou de collecter les déchets ensemble et d'effectuer le tri ensuite dans un centre de tri. Le tri sélectif nécessite la participation de chacun, du citoyen à la collectivité locale. Le tri contribue ainsi à la préservation des ressources naturelles, évite le gaspillage et réduit les coûts de traitement en limitant les quantités à incinérer. Le tri sélectif est un pas vers une écologie plus saine [23]. Pour mettre en œuvre efficacement ces différents modes de traitement, il est indispensable de mettre en place en amont un tri des déchets pour orienter chaque déchet (ou fraction de déchet) vers la solution optimale. Le meilleur tri étant situé à la source du déchet. L'idéal est bien sûr d'éviter la production de déchets.

(d) La collecte : La collecte désigne l'ensemble des opérations qui consistent à regrouper les déchets, depuis leurs sources de production puis à les transporter jusqu'aux centres de traitement [14].

- La collecte par apport volontaire : Des conteneurs spécifiques sont installés en différents points de la ville pour récupérer les produits recyclables. Les habitants apportent leurs déchets recyclables jusqu'au conteneur. De la même façon, les déchèteries reçoivent des déchets qui ne sont pas acceptés par la collecte normale.
- La collecte séparative en porte à porte : La commune fournit à ses habitants plusieurs poubelles, qui ont des couvercles de couleurs différentes selon ce qu'elles doivent

contenir. Les habitants placent leurs déchets dans les différentes poubelles. Il peut y avoir par exemple une poubelle pour les matières organiques, une poubelle pour les emballages et une dernière poubelle pour les déchets non recyclables.

- La collecte mixte : La collecte est partagée entre apport volontaire et collecte séparative en porte à porte. Par exemple, les emballages et déchets non recyclables sont récupérés en porte à porte et le verre et le papier peuvent être apportés dans des conteneurs en ville. La collecte mixte est le plus souvent la solution retenue.
 - La collecte par point de regroupement : Est particulièrement développée en milieu rural dispersé ou groupé. Ce sont les usagers qui déplacent leurs déchets ou récipients afin de les grouper avec ceux des habitants voisins. Ce système permet de diminuer le nombre de points et le temps de collecte et donc le coût supporté par la collectivité.
 - La collecte sélective : Vise la collecte de certains flux de déchets préalablement séparés par les producteurs, en vue d'une valorisation ou d'un traitement spécifique.
- (e) Le transit et le transport des déchets : La station de transit ou centre de transfert reçoit les flux de déchets ménagers et permet leur stockage dans une fosse, des conteneurs ou sur une aire. Les déchets, éventuellement compactés sont ensuite acheminés, par un mode de transport vers une installation de traitement ou de stockage. L'emploi d'une station de transit n'est justifié que dans le cas où le traitement des déchets n'est pas réalisé dans l'aire géographique de la collecte.
- (f) Le traitement : Ensemble de procédés visant à transformer les déchets pour notamment en réduire dans des conditions contrôlées leur potentiel polluant initial, la quantité ou le volume, et assurer leur recyclage ou leur valorisation. (Incinération, compostage,...).

1.5 Stratégies de la gestion des déchets :

Face à la nécessité de résoudre le problème de la gestion d'un déchet, nous allons présenter les objectifs de la gestion des déchets, ensuite les choix stratégiques proposés par [22].

1.5.1 Les objectifs de la gestion des déchets [14]

- (a) De prévenir ou de réduire la production et la nocivité des déchets, de la fabrication à la distribution des produits.
- (b) D'organiser le transport des déchets et de le limiter en distance et en volume.
- (c) De valoriser les déchets par réemploi, recyclage ou toute autre action visant à obtenir des matériaux réutilisables ou de l'énergie.
- (d) De planifier la gestion des déchets. Pour que les déchets soient transportés et éliminés conformément à ces objectifs, leur gestion fait l'objet d'une planification.

- (e) D'assurer l'information du public sur les effets pour l'environnement et la santé publique des opérations de production et d'élimination des déchets, et sur les mesures destinées à prévenir ou compenser les effets préjudiciables de ces opérations.

1.5.2 Les choix stratégiques

- (a) Arrêt de la production ou de la diffusion du produit à l'origine de déchet : cette stratégie est imposé par les impacts écologiques ou toxiques liés à l'usage de certains produits.
- (b) Optimisation des procédés et innovations technologiques : Technologies propres, sobres et économes. Le développement des biotechnologies et s'inscrit dans cette évolution des outils de production moins générateurs de déchets dangereux.
- (c) Valorisation des déchets : La majeure partie des déchets sont potentiellement valorisables en fonction de leur nature chimique, de leurs propriétés mécaniques, physico-chimiques ou thermiques (valorisation du verre, papier, plastique).
- (d) Rejet éco compatible dans le milieu naturel : Dans certains cas, le retour des déchets en milieu naturel est sans impact écologique, écotoxique ou toxique. C'est le cas en général des déchets inertes.
- (e) Stockage dans le milieu naturel : Il s'agit de l'enfouissement des déchets. C'est la mise en décharge dans un contexte réglementé qui doit garantir l'innocuité du système vis-à-vis du milieu environnant.

1.6 Impacts des déchets solides sur la santé et l'environnement[22]

L'utilisation intensive et abusive des ressources et le rejet des déchets dans l'environnement contribuent à détériorer notre milieu. Ce changement a un impact sur la société, la santé humaine, et l'environnement.

1.6.1 Impacts sur la santé humaine

- (a) **Pathologies liées à des conditions environnementales favorables et maladies spécifiques de la manipulation des déchets (agents de nettoyage, chiffonniers...)**
 - Hépatites épidémiques et sériques.
 - Conjonctivites épidémiques.
 - Tétanos.
 - Proéminence de la tuberculose.
 - Effets multiples des substances radioactives.

- Intoxications aux produits dangereux.
- Maladies de contact de la peau et des muqueuses.

(b) Impacts sanitaires des décharges non contrôlées

- Multiplication des maladies infectieuses et parasitaires (hépatites infectieuses, maladies parasitaires de la peau et autres).
- Multiplication des rongeurs qui sont à l'origine de la peste.
- Prolifération des chiens errants (zoonoses (la rage) et parasitoses (maladies liées à la tique des chiens).
- Prolifération des vecteurs nuisibles (mouches, moustiques,...).

1.6.2 Impacts sur l'environnement :

Les déchets solides ont un impact environnemental sévère qui se manifeste par une

- (a) Altération de la qualité de l'air (gaz, fumées et poussières).
- (b) Altération des sols et des paysages par des polluants chimiques.
- (c) Pollution des ressources en eau par les infiltrats et les eaux usées.

1.7 La gestion des déchets solides en Algérie

Pendant longtemps, le gestionnaire de déchets n'avait d'autre choix que celui de les mettre en décharge ou de les incinérer, de les éloigner ou de les brûler, mais surtout s'en débarrasser, sans trop s'occuper des conséquences qui, à terme, pouvaient survenir. Nous pensons que l'époque appelle aujourd'hui un changement d'attitude. Plus positive, plus économe, plus responsable. La gestion des déchets solides est une priorité environnementale et économique importante dans tous les pays. Le secteur des déchets et l'amélioration de la qualité de vie constituent actuellement un axe de travail prioritaire des autorités. Le gouvernement algérien, conscient que le secteur des déchets solides recèle de nombreuses opportunités pouvant engager son développement sur une base durable, a montré sa volonté politique pour doter le pays d'un système de gestion sain et performant [57].

1.7.1 Principes de la gestion de déchets en Algérie [30]

S'agissant de la gestion des déchets municipaux, le programme national pour la gestion intégrée des déchets ménagers PROGDEM, a été mis en œuvre par le Ministère de l'Aménagement du Territoire et de l'Environnement en 2001. Il a pour objet l'implémentation d'une nouvelle stratégie qui repose sur une série de principes ; le principe de précaution-prévention, le principe du pollueur-payeur, le principe de producteur- récupérateur, la promotion de l'information et de la sensibilisation du citoyen. Ce programme fixe des objectifs en matière d'amélioration du

cadre de vie du citoyen, de la préservation de la salubrité publique et la protection de la santé, ainsi que l'élimination saine et écologiquement rationnelle des déchets par la valorisation des déchets recyclables. Il a touché dans un premier temps les 40 grandes villes algériennes.

Les principales actions concernées par PROGDEM sont les suivantes :

- (a) L'élaboration et la mise en œuvre des plans communaux de gestion des déchets, qui sont des instruments de planification et de gestion qui permettent de projeter sur le moyen et le long terme. En effet, les études relatives à ces plans prescrivent les conditions de présentation des déchets à la collecte, fixent les normes de ramassage et l'évacuation des déchets et elles établissent les cahiers des charges qui précisent les obligations auxquelles doivent être soumises les entreprises chargées du ramassage.
- (b) L'aménagement de sites de mise en décharge contrôlée pour réduire l'impact sur l'environnement de la gestion chaotique de la filière élimination des déchets. À ce propos des centres de stockage de déchets ont été réalisés ou sont en cours de réalisation à travers toutes les wilayas du pays. Certains sont déjà opérationnels comme celui du CET d'Ouled Fayet à Alger, le Centre d'Enfouissement Technique CET d'El Outaya à Biskra. Agence Nationale des Déchets, donne le chiffre de 32 CET achevés à travers le territoire national, 42 sont en cours de réalisation, 7 autres en voie de lancement, 12 en phase d'étude et 4 centres sont en phase de choix de site. À titre d'exemple dans la wilaya d'Alger et en vue d'absorber les déchets ménagers, 4 CET sont en cours de lancement au niveau des communes de Corso, de Réghaïa, de Staouali, et d'El Hamiz. L'objectif étant d'atteindre 300 CET exploités en 2014 [46].
- (c) La promotion des activités de recyclage et de valorisation des déchets. Les responsables du MATET par exemple, considèrent que l'échec relatif des expériences précédentes en matière de valorisation particulièrement dans les unités de compostage de Badjarah, Blida et Tlemcen était davantage imputable au mélange des déchets qu'aux techniques elles-mêmes. Pourtant, selon l'Agence de l'environnement et de maîtrise d'énergie ADEM, l'absence de tri sélectif ne peut pas constituer un frein réel, aussi les déchets en Algérie peuvent être recyclés car 30% des déchets en sont recyclables. Par contre, d'après l'ADEM, pour inciter les agriculteurs locaux à épandre le compost obtenu, il est nécessaire que le Ministère de l'Agriculture donne une impulsion pour mettre en place une filière de valorisation du compost [11].
- (d) L'introduction de nouvelles formes de gestion, s'agissant par exemple de la collecte et la gestion des déchets ménagers de la wilaya d'Alger, c'est l'EPIC NET COM qui en a la charge. Créée en 1996, NET COM opère sous la tutelle de la Direction de l'Environnement. À ce jour, l'EPIC est considérée comme une forme transitoire avant d'avoir recours à d'autres formes de gestion. La gestion mixte ou la gestion déléguée ne sont pas écartées par les autorités algériennes. D'ailleurs, plusieurs solutions sont actuellement testées dans différentes communes comme les communes de Bordj El Kiffan et Kouba qui ont été

données en concession à un opérateur privé [11].

- (e) L'adaptation graduelle de la taxe d'enlèvement des déchets ménagers et l'amélioration de son taux de recouvrement; Assemblée Populaire Communale fixe le montant de la taxe (TEOM) prévue par la loi de finances de 2002.
- (f) Sensibilisation, formation et éducation à travers la création de Centre National de Formation en Environnement pour la formation des cadres d'état dans le domaine de l'environnement, et l'implantation d'une maison de l'environnement dans chaque wilaya pour la sensibilisation des citoyens et l'éducation des collégiens.

1.8 Revue sur quelques travaux connexes

Les exigences croissantes en matière de gestion des déchets solides sont devenues un défi majeur pour les autorités municipales. Un certain nombre de systèmes et de méthodes intégrés ont été introduits pour surmonter ce problème.

- (a) **Classification d'images de déchets solide avec l'algorithme de déformation temporelle dynamique et un classificateur multicouches :**

L'approche proposée par [45], consiste à créer une solution de traitement d'image visant à résoudre la problématique de la difficulté de positionner la caméra pour obtenir une image centralisée de la zone de la corbeille. L'algorithme de déformation temporelle dynamique a été utilisé pour détecter et recadrer la zone de la corbeille, tandis que les ondelettes de Gabor ont été introduites pour l'extraction des caractéristiques de l'image de la corbeille. Les caractéristiques d'image ont été utilisées pour l'entraînement du classificateur. Un classificateur perceptron multicouches a été utilisé pour classifier le niveau de la poubelle et estimer la quantité de déchets à l'intérieur de la poubelle. La zone située sous les courbes ROC (Receiver Operating Characteristic) a été utilisée pour évaluer statistiquement les performances du classifieur. La démonstration du système a donné des résultats de 98.5% de précision pour l'estimation du niveau de déchet.

- (b) **Système de récupération d'images basé sur le contenu pour la détection du niveau de la corbeille à déchets et évaluation des performances :**

Selon [31], une méthodologie a été proposée pour la création d'un système pour la récupération d'une texture extraite de l'image d'une corbeille pour détecter le niveau de corbeille. Afin d'obtenir les performances les plus élevées, différentes distances de similarité telles que Euclidean, Bhattacharyya, Chi-carré, Cosinus et EMD ont été utilisées pour calculer et comparer la distance entre une image de requête et les images d'une base de données. Dans cette étude, les mesures de performance reposent sur deux critères d'évaluation quantitatifs. Le premier est le taux de récupération moyen basé sur le graphe de précision-rappel et le second est l'utilisation de la mesure F1. En cas d'extraction de caractéristiques, la texture est utilisée comme caractéristique d'image pour le système de

détection de niveau de déchet. Diverses expériences ont été effectuées avec différentes techniques d'extraction de caractéristiques, telles que le filtre à ondelettes de Gabor, GLCM et la matrice d'Aura de niveaux de gris (GLAM) afin d'identifier le niveau de la corbeille et la zone qui l'entoure. Les tests et la validation de la classification ont été effectués par les classificateurs Perceptron multicouche et K-ppv. Le système développé avait montré des bons résultats avec un taux de précision de 95% en utilisant le classificateur Perceptron multicouche, tandis que 97% en utilisant le classificateur K-ppv.

- (c) **Un modèle mathématique pour la gestion des déchets solides municipaux - Une étude de cas à Hong Kong :** Un modèle mathématique a été créé par [40], ce dernier adopte une programmation linéaire en nombres entiers, et une programmation en nombres entiers mixtes a été mise au point pour la gestion des déchets solides municipaux de Hong Kong. Une analyse de sensibilité a été réalisée pour simuler différents scénarios, fournissant aux décideurs des informations importantes pour la mise en place d'un système de gestion des déchets. Les deux approches ont révélé des résultats similaires, mais les résultats du modèle à nombres entiers mixtes (seconde) sont relativement meilleurs, car le coût est basé sur le coût unitaire réel plutôt que sur le coût estimé calculé sur la capacité pleine de chargement du camion.
- (d) **Un système automatisé de détection du niveau des déchets solides utilisant des filtres à ondelettes de Gabor et une perception multicouche :** Une autre méthode a été proposée par [32], qui utilise un filtre à ondelettes de Gabor comme extracteur de caractéristiques. Profitant des caractéristiques souhaitables des filtres de Gabor, telles que la taille du masque et la sélectivité en fréquence, ils ont conçu quatre filtres correspondant à différentes valeurs de fréquence pour l'extraction de caractéristiques des déchets à partir d'images. La sortie du vecteur de caractéristiques basée sur les filtres de Gabor est utilisée comme entrée de l'algorithme de classification, qui est un réseau de neurones à propagation. Le système a été testé sur un grand nombre d'images et a obtenu une classification de 94% de précision .

1.9 Conclusion

Dans ce chapitre, nous avons introduit les définitions et les stratégies liées au domaine de la gestion des déchets solides. Nous avons présenté et discuté les solutions existantes. Dans le chapitre suivant, nous présentons les différentes techniques de classification d'images.

Chapitre 2

Techniques de classification d'images

La classification d'images basée sur le contenu «Content-Based Image Retrieval (CBIR)» utilise des caractéristiques visuelles pour rechercher et récupérer des images numériques à partir d'une base de données volumineuse. Les performances d'un système de classification d'images basés sur le contenu dépendent de manière cruciale de la représentation des caractéristiques. A partir d'un ensemble d'images d'apprentissage, l'objectif du système est d'inférer un lien entre les caractéristiques de chaque image et sa classe. Dans la littérature, plusieurs méthodes ont été proposées pour résoudre le problème de classification d'image en se basant sur le contenu, ces méthodes sont dites classiques, d'autres adaptatives [33].

La suite de ce chapitre est organisée en deux sections :

- **La première section** : présente les concepts généraux des méthodes, dites, classiques de la recherche d'images par contenu, ainsi que l'extraction des caractéristiques visuelles.
- **La deuxième section** est consacrée à la présentation des concepts des méthodes adaptatives basées sur l'apprentissage supervisé. Nous nous intéressons plus particulièrement aux réseaux de neurones, notamment les réseaux de neurones convolutifs. Et enfin nous présentons la technique d'apprentissage par transfert et ses différentes stratégies pour la classification d'images.

2.1 Système d'indexation et de recherche d'images

Les systèmes de Recherche d'images basés sur le contenu « Content-Based Image Retrieval (CBIR)» ont émergé au début des années 1990 [21]. Ils permettaient d'assigner des signatures aux images à base de leurs caractéristiques visuelles, afin de faciliter leur récupération par la suite. Dans ce qui suit, nous allons présenter le concept général de la recherche d'image et l'architecture général d'un système de recherche.

2.1.1 Principe général de l'indexation de la recherche d'information

L'indexation d'un document dans une base de données consiste à lui attribuer une signature qui va faciliter de le retrouver plus tard. L'utilisateur fournit un document requête et le système recherche les documents les plus similaires. En général, les systèmes de recherches comprennent les éléments suivants :

- (a) **Une signature (ou index du document)** qui sert comme caractéristique pour le reconnaître et le comparer avec les autres. Dans le cas de la recherche par caractéristiques, celles-ci ou les dérivées de ces caractéristiques sont prises directement comme signatures. Quant à la recherche par l'exemple, il faut déterminer les signatures les plus appropriées pour décrire le contenu des documents d'une façon approchant le mieux possible les critères de l'utilisateur.
- (b) **Une métrique de similitude (ou de distance)** qui permet de comparer les signatures et d'associer les documents similaires.
- (c) **Des algorithmes de recherche** qui sont basés sur les deux outils précédents. Ils permettent de retrouver rapidement les objets recherchés. Une approche itérative est parfois proposée : à partir d'une première sélection de documents, nous précisons les critères, nous relançons la requête, et ainsi de suite.
- (d) **Une interface utilisateur** qui rend la procédure de recherche transparente et facilite l'introduction de la requête [49].

La figure 2.1 illustre le processus d'un système de recherche d'information RI

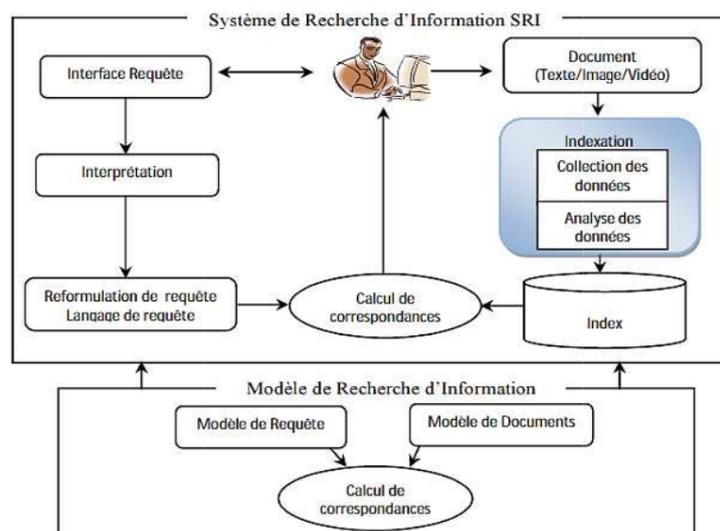


FIGURE 2.1 – Processus de RI

[56]

La recherche d'images par le contenu -CBIR- diffère de la recherche d'information textuelles par le fait que les images ne contiennent pas des informations sémantiques interprétable par les êtres humain. L'image est un ensemble de pixels représentés par une matrice d'intensités sans signification.

2.1.2 Concepts fondamentaux de la recherche d'image :

(a) Définition d'une image

Une image est une représentation visuelle de divers sujets (objets, personnes, phénomène naturel, monuments, etc.). Les premières images sont apparues dans la période préhistorique sous la forme de dessins destinés à communiquer des messages. Au fil du temps, le concept d'image a évolué : de la période du Moyen-âge et la renaissance sous forme de tableau, jusqu'à nos jours sous forme d'images numériques[47]. Les images numériques représentent des images qui peuvent être stockées sur un support informatique (ordinateur, téléphone, tablette, clé USB. . .) sous forme d'une matrice de Pixel (abréviation de la locution anglaise « picture element») [47].

Une image numérique est définie par :

- le nombre de pixels qui la composent en largeur et en hauteur.
- l'étendue des teintes de gris ou des couleurs que peut prendre chaque Pixel.

Nous distinguons trois types d'images numériques [44] :

- **Les images binaire** : (noire et blanc), ce sont les images les plus simple, le pixel ne peut prendre qu'à la valeur blanche ou noir.
- **Les images aux niveaux de gris** : En général, les images aux niveaux de gris renferment 256 teintes de gris. Par convention la valeur 0 représente le noir (intensité lumineuse nulle) et la valeur 255 le blanc (intensité lumineuse maximale). Le nombre 256 est lié à **la quantification** de l'image. En effet chaque entier représentant un niveau de gris est codé sur 8 bits. Il est donc compris entre 0 et $2^8 - 1 = 255$. C'est la quantification la plus courante. Nous pouvons coder une image aux niveaux de gris sur 16 bits ($0 \leq n \leq 2^{16} - 1$) ou sur 2 bits.
- **Les images en RGB** : L'espace couleur est basé sur la synthèse additive des couleurs. Dans ce type d'images un pixel est codé par trois valeurs numériques. La signification de ces valeurs dépend du type de codage choisi.

(b) Définition de la recherche d'images :

Par analogie de la recherche d'information textuelle à la recherche d'images avec le domaine de la RI, les images peuvent être considérées comme des documents et un système de recherche d'images est défini comme « le système qui s'intéresse à la navigation, la recherche et l'extraction d'images à partir d'une grande base de données d'images numériques » [47].

(c) **Les principales approches pour l'indexation d'images :**

Il existe plusieurs techniques d'indexation d'images proposées dans la littérature; commençant par la première technique qui consiste à décrire le contenu visuel par des mots clés . Ces mots-clés servent comme index pour accéder aux données visuelles associées. L'avantage de cette approche est qu'elle permet de consulter les bases de données en utilisant les langages d'interrogation standard, tels que SQL. Mais elle nécessite une grande quantité de traitement manuel. De plus, la description des données est subjective et dépendante de la langue utilisée (anglais, français, espagnole, etc...). Afin de pallier ces lacunes, les systèmes d'indexation par le contenu sont introduits. Ces systèmes sont basés sur l'extraction de caractéristiques numériques, où les index servent pour encoder le contenu des images. Nous nous intéressons plus particulièrement à l'approche basée sur l'extraction de caractéristiques numériques. Nous développons l'architecture des systèmes d'indexation par contenu au paragraphe suivant [56].

2.1.3 Architecture des systèmes d'indexation par contenu

Un système d'indexation par contenu se compose de deux principales phases, la première, en mode hors ligne, elle consiste à faire les prétraitements, la gestion et l'organisation de la base de données de références, autrement dit le sous système d'indexation. La deuxième étape qui se fait en ligne, est la phase de traitement à l'arrivée d'une nouvelle requête. Pour les deux sous systèmes, nous devons passer par l'extraction des descripteurs caractéristiques des images lors de la phase en ligne et à base de ces caractéristiques extraites nous effectuons les index. L'indexation des images consiste à calculer les signatures des images à base des descripteurs. Plus deux signatures sont proches au sens de la métrique de similarité utilisée, plus les images associées sont proches au sens de l'utilisateur. La complexité de la comparaison des images est alors réduite à la comparaison de leurs signatures d'indexation pendant la consultation de la base, Finalement, le système sélectionne et présente à l'utilisateur les images les plus similaires à la requête [49].

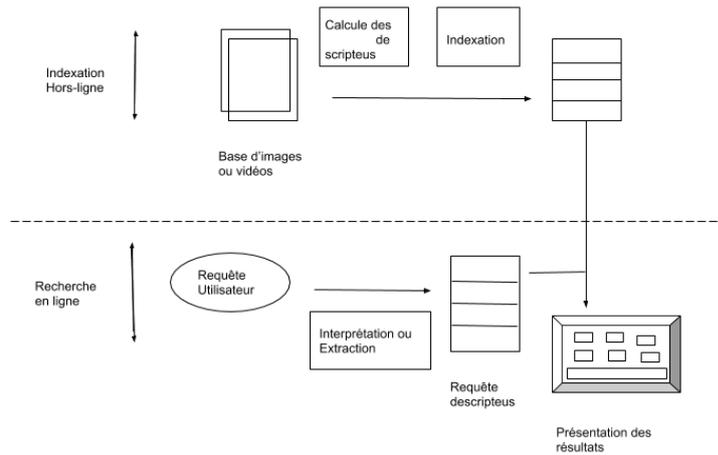


FIGURE 2.2 – Architecture d’un système d’indexation et de recherche d’images

[13]

2.1.4 Principales techniques d’indexation par contenu

Les techniques d’indexation d’images exploitent plusieurs types de caractéristiques de bas niveau, dites visuelles, tel que :

- (a) **La couleur** : est considérée comme l’une des caractéristiques les plus importantes et les plus utilisées dans la recherche d’images par contenu [56]. Malgré son importance, se baser seulement sur elle n’est pas suffisant puisque dans le cas de deux images dont l’histogramme de couleur est identique, leurs concepts sont différents [47].
- (b) **La texture** : est le second attribut visuel largement utilisé dans la recherche d’images par le contenu, elle est définie comme “la répétition d’un motif créant une image visuellement homogène. Plus précisément, la texture peut être vue comme un ensemble de pixels (niveaux de gris) spatialement agencés selon un certain nombre de relations spatiales, ainsi créant une région homogène” [28]. De nombreuses approches et modèles ont utilisé la texture pour l’indexation [47], parmi les plus connus nous pouvons citer : les méthodes statistiques. Nous introduisons dans le paragraphe suivant la représentation des caractéristiques en utilisant la matrice de co-occurrence des niveaux de gris.

— **Extraction des caractéristiques en utilisant la matrice de co-occurrence des niveaux de gris** : La matrice de co-occurrence des niveaux de gris ou bien “Gray Level Co-occurrence Matrix “ en anglais - sous l’acronyme “GLCM” - est un moyen de caractérisation d’une image basé sur la texture.

GLCM consiste à étudier la texture dans une image en comptabilisant le nombre de paires de Pixel distant de d qui représente la différence δz entre les niveaux de gris. En plus de la

distance, on tient compte de la direction θ définie par toute paire de pixels[13].

Nous pouvons définir la matrice de co-occurrence par une relation R entre deux pixels (X_1, Y_1) et (X_2, Y_2) , un simple exemple d'une relation géométrique est le suivant :

$$X_2 = X_1 + 2$$

$$Y_2 = Y_1$$

pour lequel le Pixel (X_2, Y_2) est le Pixel voisin de droite de (X_1, Y_1) . La matrice de co-occurrence $C(i, j)$ est une matrice carrée de taille $L \times L$ où L est le nombre de niveaux de gris présent, les indices de la matrice sont les niveaux de gris de la texture étudiée; la matrice de co-occurrence se construit comme suit :

Initialiser la matrice : $\forall i, j \in [0, l[: C(i, j) = 0$

Remplir la matrice : vérifier si la relation R est vérifiée, si c'est le cas :

$$C_R(f(X_1, Y_1), f(X_2, Y_2)) = C_R(f(X_1, Y_1), f(X_2, Y_2)) + 1$$

À partir du matrice de co-occurrence, nous pouvons effectuer plusieurs statistiques qui fournissent des informations sur la texture de l'image. Nous citons les statistiques suivantes :

- **Contraste** : Il mesure les variations locales dans la matrice de co-occurrence de niveaux gris [31].

$$Contraste = \sum_{i,j} |i - j|^2 p(i, j) \quad (2.1)$$

- **Corrélation** : mesure l'occurrence de probabilité conjointe des paires de Pixel spécifiées[31].

$$Correlation = \sum_{i,j} \frac{(i - \mu_i)(j - \mu_j)p(i, j)}{\sigma_i \sigma_j} \quad (2.2)$$

- **Énergie** : Elle fournit la somme des éléments carrés dans la GLCM. Aussi connu comme l'uniformité ou le second moment angulaire [31].

$$Energie = \sum_{i,j} p(i, j)^2 \quad (2.3)$$

- **Uniformité** : Elle mesure la proximité de la distribution des éléments dans la GLCM à la diagonale GLCM [31].

$$Uniformite = \sum_{i,j} \frac{p(i, j)}{1 + |i - j|} \quad (2.4)$$

- **Entropie** : Une mesure de l'aléatoire que nous utilisons pour caractériser l'image [43]

$$Entropie = \sum_i p \times \log_2(p) \quad (2.5)$$

2.1.5 Les mesures de similarité :

Plusieurs méthodes proposent de mesurer la similitude entre deux images , ces méthodes consistent à calculer une distance entre les deux vecteurs de caractéristiques, nous montrons les mesures de similarité suivantes :

— La distance euclidienne [31] :

$$D_{euclid} = \sqrt{\sum_{i=1}^d (x_i - y_i)^2} \quad (2.6)$$

— La distance cosinus [31] :

$$D_{cosine} = \frac{\sum_{i=1}^d x_i y_i}{\sqrt{\sum_{i=1}^d x_i^2} \sqrt{\sum_{i=1}^d y_i^2}} \quad (2.7)$$

— La corrélation [42] :

$$D_{correlation} = 1 - \frac{(x_s - \bar{x}_s)(y_t - \bar{y}_t)'}{\sqrt{(x_s - \bar{x}_s)(x_s - \bar{x}_s)'}} \sqrt{(y_t - \bar{y}_t)(y_t - \bar{y}_t)'}} \quad (2.8)$$

— distance Bhattacharyya :

$$D_{batt}(x, y) = \frac{1}{8}(M_x - M_y)^T \left[\frac{C_x + C_y}{2} \right] (M_x - M_y) + \frac{1}{2} \ln \frac{|\frac{M_x + M_y}{2}|}{\sqrt{|C_x||C_y|}} \quad (2.9)$$

2.2 Techniques d'apprentissage automatique :

L'apprentissage automatique (machine learning en anglais), qui est l'un des sous-domaines de l'intelligence artificielle, a pour objectif d'extraire et d'exploiter automatiquement l'information présente dans un jeu de données.

Bien que l'apprentissage automatique soit un sous domaine de l'informatique, il diffère des approches informatiques traditionnelles ; il facilite l'utilisation des ordinateurs dans la construction de modèles à partir de données d'échantillonnage afin d'automatiser les processus de prise de décision en fonction des données saisies [50].

Les algorithmes d'apprentissage peuvent se catégoriser selon le type d'apprentissage qu'ils emploient [12] :

2.2.1 Apprentissage supervisé :

Dans le contexte de l'intelligence artificielle et l'apprentissage automatique, l'approche supervisée est un système qui fournit un modèle de donnée déjà étiqueté par un ou plusieurs experts de domaine, ce modèle est utilisé pour le traitement ultérieur de nouvelle donnée [47]. Il existe plusieurs approches d'apprentissage supervisé -nous mentionnons les plus employées- :

- (a) **Les K-plus proches voisins (Kppv)** : est un algorithme très populaire, dédié aux problèmes de classification, son principe est très simple, il s'agit de trouver les images : x_1, x_2, \dots, x_k les plus proches de la requête x à partir du modèle de données étiquetées, puis attribuer la classe correspondante à x selon le résultat majoritaire des statistiques de classe d'appartenance [17].
- (b) **Machines à Vecteurs de Support (SVM)** : est une méthode de classification statistique proposée par Vapnik [8]. qui se base sur le principe structurel de minimisation des risques à partir de la théorie de l'apprentissage informatique. SVM vise à trouver une surface de décision pour séparer les points de données d'entraînement en deux classes et prend des décisions sur la base des vecteurs de support qui sont sélectionnés comme les seuls éléments efficaces dans l'ensemble de la formation [35].

2.2.2 Apprentissage non-supervisé :

Contrairement à l'apprentissage supervisé, l'apprentissage non supervisé consiste à cibler les groupes homogènes d'exemples existants dans les données qui sont à la base non étiquetés [15]. Contrairement aux approches supervisées, les approches non-supervisées ont besoin d'un nombre plus grand de données pour le processus d'apprentissage [47].

Parmi les algorithmes mis en œuvre dans l'apprentissage non supervisé, nous mentionnons K-moyennes, k-médoïdes, etc.

2.2.3 Apprentissage profond

L'apprentissage profond fait référence à un type d'Intelligence Artificielle particulier utilisant notamment le réseau de neurones et de certains modèles d'algorithmes particuliers comme le modèle de neurone convolutif (convolutional neural network) afin de générer des modèles intelligents grâce à l'apprentissage.

(a) Les réseaux de neurones

- **Perceptron** : Les perceptrons sont des classificateurs binaires, inventés par [52]. le développement des perceptrons a été inspiré par l'architecture du cerveau humain, ce dernier est composé de cellules appelées neurones qui traitent les informations, les connexions entre neurones appelées synapses par lesquelles les informations sont transmises. Les chercheurs estiment que le cerveau humain est composé de 100 milliards de neurones et de 100 trillions de synapses. Les principaux composants d'un neurone sont des dendrites, un corps et un axone. Les dendrites reçoivent des signaux électriques d'autres neurones. Les signaux sont traités dans le corps du neurone, qui envoie ensuite un signal via l'axone à un autre neurone [29]. un neurone individuel, peut être vu comme une unité de calcul, qui traite une ou plusieurs entrées, et produit

par la suite une sortie. Les Perceptrons sont généralement visualisés à l'aide d'un diagramme comme celui-ci(2.3) :

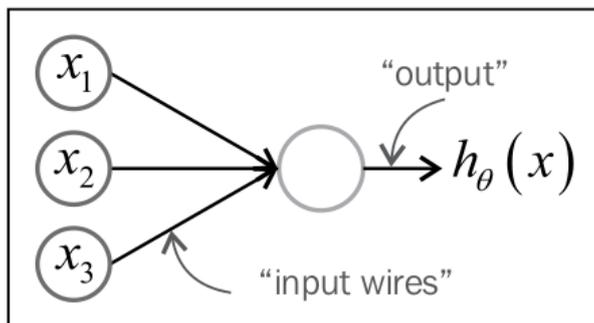


FIGURE 2.3 – L'architecture générale d'un Perceptron [29].

Les cercles x_1 , x_2 et x_3 sont des unités en entrée. Chaque unité d'entrée représente une caractéristique. Les perceptrons utilisent fréquemment une unité d'entrée supplémentaire représentant un terme de biais constant, mais cette unité d'entrée est généralement omise des diagrammes. Le cercle au centre est une unité de calcul ou le corps du neurone. Les arêtes reliant les unités d'entrée à l'unité de calcul sont analogues aux dendrites. Chaque arête est pondérée ou associée à un paramètre(poids). Les paramètres peuvent être interprétés facilement. une variable explicative corrélée à la classe positive aura une pondération positive et une variable explicative corrélée à la classe négative aura une pondération négative. Le bord éloigné de l'unité de calcul renvoie la sortie.

- **Fonction d'activation** : Le Perceptron classe les instances en traitant une combinaison linéaire des variables explicatives et des paramètres du modèle à l'aide d'une fonction d'activation, comme indiqué dans l'équation suivante. La combinaison linéaire des paramètres et des entrées est parfois appelée préactivation du perceptron.

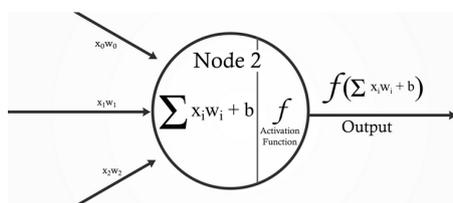


FIGURE 2.4 – Schéma démonstratif de la fonction d'activation [58].

Ici, w_i sont les paramètres du modèle, b est un terme de biais constant et F est la fonction d'activation. Plusieurs fonctions d'activation différentes sont couramment utilisées. Le perceptron original de Rosenblatt utilisait la fonction

de **Heaviside** . Également appelée fonction de pas unitaire, la fonction de pas Heaviside est montrée dans l'équation suivante, où x est la combinaison pondérée des caractéristiques :

$$g(x) = \begin{cases} 1, & x > 0 \\ 0, & \text{elsewhere} \end{cases} \quad [29]$$

Si la somme pondérée des variables explicatives et du terme de biais est supérieure à zéro, la fonction d'activation renvoie un et le perceptron prédit que l'instance est la classe positive. Sinon, la fonction renvoie zéro et le perceptron prédit que l'instance est la classe négative. La fonction d'activation **Heaviside** est représentée dans la figure 2.5 :

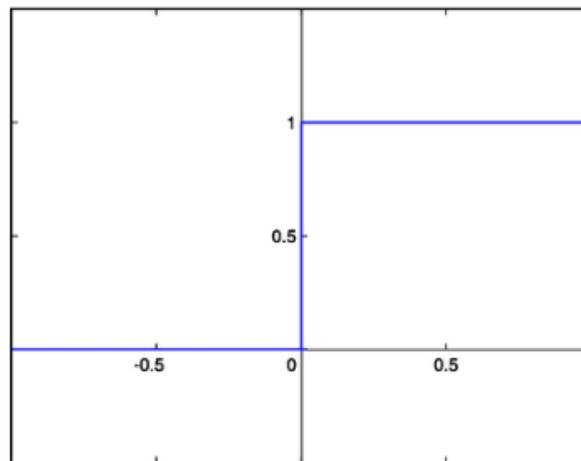


FIGURE 2.5 – Graphe de la fonction Heaviside [29].

- **Les limites de Perceptron :** Les modèles linéaires comme le perceptron avec une fonction d'activation **Heaviside** ne sont pas des **approximateurs universels de fonctions** ; ils ne peuvent pas représenter certaines fonctions. Plus précisément, les modèles linéaires ne peuvent apprendre à approximer les fonctions que pour des données linéairement séparables. Les classificateurs linéaires, trouvent un hyperplan qui sépare les classes positives des classes négatives ; s'il n'existe aucun hyperplan pouvant séparer les classes, le problème n'est pas linéairement séparable. Une des techniques pour modéliser un problème avec des données linéairement inséparable, c'est les réseaux de neurones.
- **Du perceptron aux réseaux de neurones artificiels :** les réseaux de neurones artificiels (ANN) sont des modèles puissants non-linéaires de classification et de régression, qui utilisent une stratégie différente pour surmonter les limites du perceptron. Ces réseaux utilisent plus d'une couche cachée de neurones, contrairement au perceptron à couche unique. Celles-ci sont également connues sous le nom de réseaux neuronaux profonds. Le perceptron multicouche (MLP) est l'un des réseaux

neuronaux artificiels les plus couramment utilisés. Les couches du MLP forment un graphe acyclique dirigé. Généralement, chaque couche est entièrement connectée à la couche suivante. la sortie de chaque neurone artificiel dans une couche est une entrée pour chaque neurone artificiel de la couche suivante vers la sortie. Les MLP ont au moins trois couches de neurones artificiels. La couche d'entrée est constituée de simples neurones. Les neurones d'entrée sont connectés à au moins une couche cachée de neurones artificiels. La couche cachée représente les variables latentes ; l'entrée et la sortie de cette couche ne peuvent pas être observées dans les données d'apprentissage. Enfin, la dernière couche masquée est connectée à une couche de sortie. La figure 2.6 décrit l'architecture d'un perceptron multicouche [29].

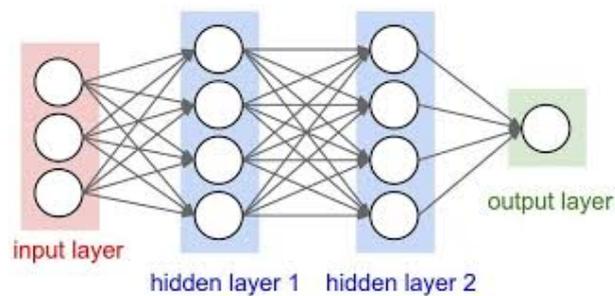
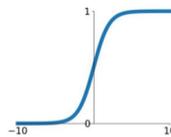


FIGURE 2.6 – L'architecture générale d'un MLP [18].

Les neurones artificiels, ou unités, de la couche cachée utilisent couramment des fonctions d'activation non linéaires, telles que la fonction de la tangente hyperbolique, la fonction logistique (sigmoid) et la fonction Rectified Linear Unit(ReLU), qui sont données par les équations suivantes :

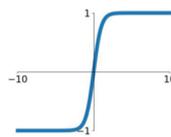
Sigmoid

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



tanh

$$\tanh(x)$$



ReLU

$$\max(0, x)$$

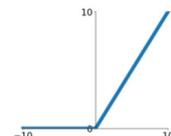


FIGURE 2.7 – Les fonctions d'activation utilisés par les réseaux de neurones [37].

— **Les paramètres d'un réseau de neurones :**

- a) **Fonction d'optimisation :** L'optimisation est un élément important dans les algorithmes d'apprentissage automatique. Il commence par définir un type de fonction de perte et finit par le minimiser à l'aide d'une procédures d'optimisation. Le choix de l'algorithme d'optimisation peut faire la différence entre obtenir une bonne précision en heures ou en jours. Parmi ces algorithmes, nous avons :

Stochastic Gradient Descent : La descente de gradient stochastique (SGD) est l'algorithme d'optimisation le plus simple permettant de trouver des paramètres minimisant la fonction de coût donnée [2].

ADAM : Cette méthode calcule les taux d'apprentissage adaptatifs pour chaque paramètre. Adam conserve une moyenne en décroissance exponentielle des gradients passés au carré et des gradients passés. Ces deux moyennes sont des estimations du premier moment (la moyenne) et du deuxième moment (la variance non centrée) des gradients [60].

RMSProp : RMSProp tente également d'atténuer les oscillations, mais d'une manière différente de la quantité de mouvement. RMSprop supprime également la nécessité d'ajuster le taux d'apprentissage et le fait automatiquement. De plus, RMSProp choisit un taux d'apprentissage différent pour chaque paramètre [2].

- b) **Fonction de perte :** Les machines apprennent au moyen d'une fonction de perte. C'est une méthode permettant d'évaluer à quel point un algorithme spécifique modélise les données fournies. Si les prévisions s'écartent trop des résultats réels, la fonction de perte en couvrirait un très grand nombre. Progressivement, à l'aide d'une fonction d'optimisation, la fonction de perte apprend à réduire l'erreur de prévision. Il existe plusieurs fonctions de perte, entre autre " categorical cross-entropy" que nous avons implémentés dans notre modèle.

L'entropie croisée (categorical cross-entropy) : La perte d'entropie croisée, ou perte logarithmique, mesure la performance du modèle de classification dont la sortie est une probabilité comprise entre 0 et 1. L'entropie croisée augmente à mesure que la probabilité prédite d'un échantillon diverge de la valeur réelle. Par conséquent, prévoir une probabilité de 0,05 lorsque l'étiquette réelle a la valeur 1 augmente la perte d'entropie croisée [4].

Mathématiquement, elle est donnée par l'équation suivante :

$$CE = - \sum_x p(x) \log q(x) \quad (2.10)$$

- c) **Le taux d'apprentissage :** Le taux d'apprentissage est probablement l'aspect le plus important de la descente de gradient, ainsi que d'autres optimiseurs. L'hyperparamètre de taux d'apprentissage contrôle le taux ou la vitesse d'apprentissage

du modèle. Plus précisément, il contrôle la quantité d'erreurs corrigées lors de la mise à jour des poids du modèle, comme à la fin de chaque lot de données d'apprentissage. Il n'y a pas de bonne valeur générique pour le taux d'apprentissage. Cela se résume à l'expérimentation et à l'intuition [38].

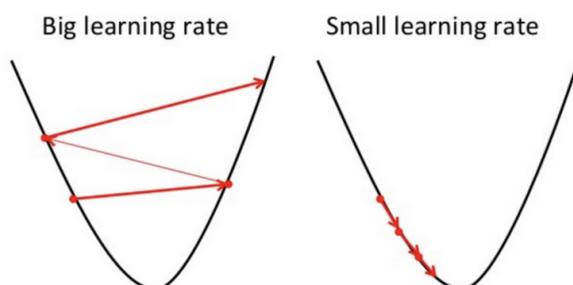


FIGURE 2.8 – Différents taux d'apprentissage [51].

(b) Les réseaux de neurones convolutifs

La structure d'un réseau de neurones convolutifs CNN a été proposée pour la première fois par Fukushima en 1988 [26]. Il n'a cependant pas été largement utilisé en raison des limites imposées par le matériel informatique nécessaire à l'entraînement du réseau. Dans les années 1990, LeCun et al. ont appliqué un algorithme d'apprentissage basé sur un gradient aux CNN et ont obtenu de bons résultats pour le problème de classification manuscrite des chiffres [39].

Les chercheurs ont ensuite amélioré les CNN et ont rapportés un état de l'art sur les résultats les plus avancés dans de nombreuses tâches de reconnaissance. Les CNN présentent plusieurs avantages par rapport aux réseaux de neurones classiques, notamment le fait d'être plus proches du système de traitement visuel humain, d'avoir une structure hautement optimisée pour le traitement des images 2D et 3D et d'être efficaces pour apprendre et extraire des abstractions de caractéristiques 2D [9].

Les réseaux CNN ont une méthodologie similaire à celle des méthodes traditionnelles d'apprentissage supervisé : ils reçoivent des images en entrée, détectent les caractéristiques de chacune d'entre elles, puis entraînent un classifieur dessus. Cependant, les caractéristiques sont apprises automatiquement, les réseaux de neurones convolutifs réalisent eux-mêmes la tâche d'extraction et de description de caractéristiques. De plus, l'architecture spécifique du réseau permet d'extraire des caractéristiques de différentes complexités, des plus simples au plus sophistiquées. L'extraction et l'hiérarchisation automatiques des caractéristiques, qui s'adaptent au problème donné, constituent une des forces des réseaux

de neurones convolutifs .

Aujourd'hui, les réseaux de neurones convolutifs, sont toujours les modèles les plus performants pour la classification d'images.

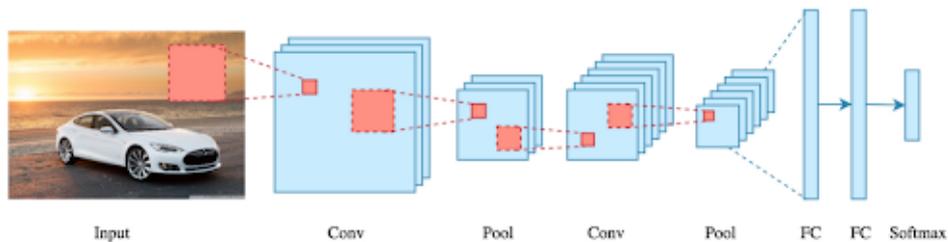


FIGURE 2.9 – L'architecture globale d'un réseau de neurones convolutifs [19].

(a) **L'architecture d'un réseau de neurones convolutifs** : Contrairement à un simple réseau de neurones, les couches d'un CNN sont rangées en 3 dimensions(largeur, hauteur, profondeur), où la profondeur est liée au type de couleurs de l'image en entrée

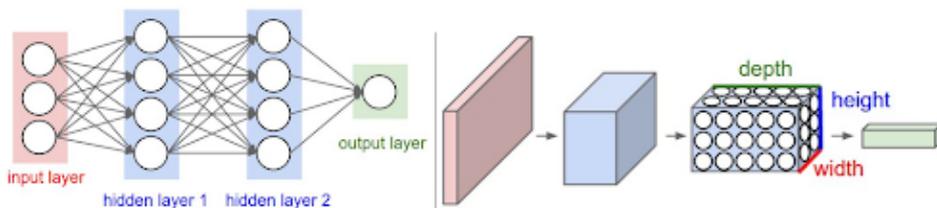


FIGURE 2.10 – un simple réseau de neurones avec 3 couches vs un CNN avec 3 couches en 3 dimensions [1].

Il existe trois types de couches pour un réseau de neurones convolutif :

- **La couche de convolution** : La couche de convolution est la composante clé des réseaux de neurones convolutifs, et constitue toujours au moins leur première couche. Son but est de repérer la présence d'un ensemble de caractéristiques dans les images reçues en entrée. Pour cela, nous réalisons un filtrage par convolution : le principe est de faire "glisser" une fenêtre représentant la caractéristique sur l'image, et de calculer le produit de convolution entre la caractéristique et chaque portion de l'image balayée. Une caractéristique est alors vue comme un filtre . La couche de convolution reçoit donc en entrée plusieurs images, et calcule la convolution de chacune d'entre elles avec chaque filtre. Les filtres correspondent exactement aux caractéristiques que l'on souhaite retrouver dans les images. Nous effectuons plusieurs convolutions sur une entrée, chacune utilisant un filtre différent et résultant en une carte de caractéristiques distincte. Nous empilons

ensuite toutes ces cartes de caractéristiques ;qui nous indique où se situent les caractéristiques dans l'image, et cela devient la sortie finale de la couche de convolution, comme le montre la figure suivante [19].

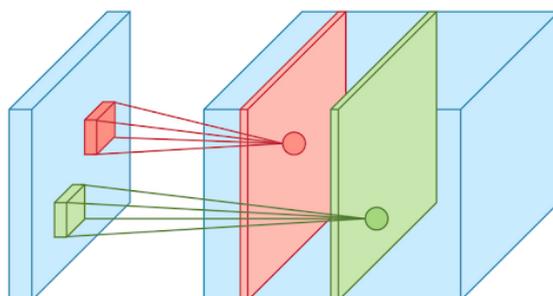


FIGURE 2.11 – Application de plusieurs filtres sur une image d'entrée [19].

- **La couche de pooling** : Ce type de couche est souvent placé entre deux couches de convolution : elle reçoit en entrée plusieurs vecteurs de caractéristiques, et applique pour chaque vecteur l'opération de pooling. L'opération de pooling consiste à réduire la taille des images, tout en préservant leurs caractéristiques importantes. Pour cela, nous découpons l'image en cellules régulières, puis nous gardons au sein de chaque cellule la valeur maximale. En pratique, nous utilisons souvent des cellules carrées de petite taille pour ne pas perdre trop d'informations. Les choix les plus communs sont des cellules adjacentes de taille 2×2 pixels qui ne se chevauchent pas, ou des cellules de taille 3×3 pixels, distantes les unes des autres d'un pas de 2 pixels (qui se chevauchent donc). On obtient en sortie le même nombre de vecteurs de caractéristiques qu'en entrée, mais ces derniers sont bien plus petits. La couche de pooling permet de réduire le nombre de paramètres et de calculs dans le réseau. Nous améliorons ainsi l'efficacité du réseau et nous évitons le sur-apprentissage. Ainsi, la couche de pooling rend le réseau moins sensible à la position des caractéristiques : le fait qu'une caractéristique se situe un peu plus en haut ou en bas, ou même qu'elle ait une orientation légèrement différente ne devrait pas provoquer un changement radical dans la classification de l'image.
- **La couche fully-connected** : Après la convolution et les couches de pooling, nous ajoutons quelques couches entièrement connectées pour envelopper l'architecture CNN. C'est la même architecture des réseaux de neurones entièrement connectée dont nous avons parlé dans la première partie.

2.2.4 Apprentissage par transfert

l'apprentissage chez l'être humain se base sur des moyens d'héritage pour transférer les connaissances entre les tâches, plus la nouvelle tâche est proche de l'expérience précédente, plus elle devient facile et maîtrisable. En revanche, la majorité des algorithmes d'apprentissage automatique traitent les tâches de manière isolée et impose la disponibilité de grande masse de données, tandis que les problèmes du monde réel possède malheureusement peu de donnée. l'apprentissage par transfert tente de changer ça en fournissant des méthodes et des techniques permettant le transfert d'une connaissance apprise dans une ou plusieurs tâches, et l'utiliser pour améliorer l'apprentissage dans les prochaines tâches [20] .

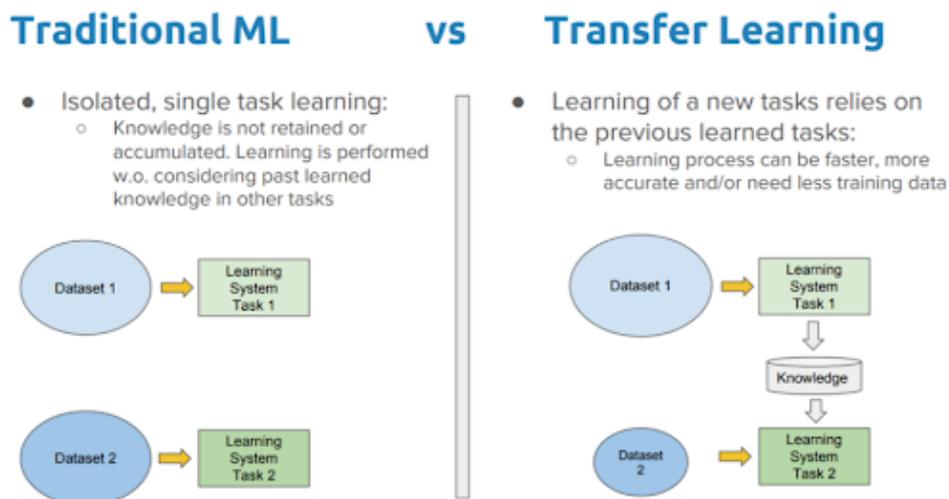


FIGURE 2.12 – l'apprentissage traditionnel et l'apprentissage par transfert [53].

Une définition globale et simple peut être “ l'apprentissage par transfert est une technique d'apprentissage automatique, où un modèle déjà entraîné sur une large base de données, sera utilisé pour une autre tâche de classification [48].

(a) **Les avantages de l'apprentissage par transfert** Le but de l'apprentissage par transfert, est d'améliorer l'apprentissage dans la tâche cible, en exploitant les connaissances de la tâche source. Cependant, selon [?] trois mesures par lesquelles nous pouvons dire que l'apprentissage par transfert, est un bon choix :

- **L'amélioration de la performance initiale** : lorsque nous augmentons les connaissances d'un apprenant isolé avec celles d'un modèle source, les performances de base peuvent être améliorées grâce à ce transfert de connaissances.
- **Durée de développement du modèle** : L'utilisation des connaissances d'un modèle source peut également aider à apprendre pleinement la tâche cible, par rapport à un modèle cible qui apprend à partir de zéro. Ceci, à son tour, entraîne des améliorations du temps global nécessaire pour développer et apprendre un modèle.

- **L'amélioration de la performance finale** : Des performances finales supérieures peuvent être atteintes en optimisant l'apprentissage par transfert.

La figure suivante, illustre les meilleures performances qui peuvent être atteintes par l'apprentissage par transfert .

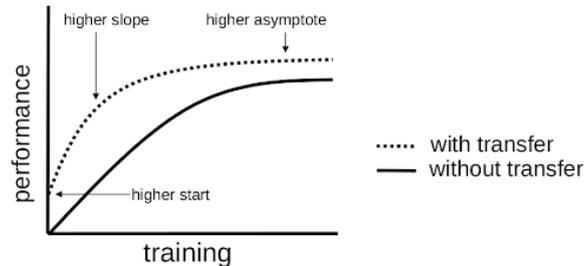


FIGURE 2.13 – Les trois mesures qui peuvent améliorer l'apprentissage par le transfert [59].

- (b) **Les stratégies de l'apprentissage par transfert** : Entraîner un réseau de neurones convolutifs à zéro, exige la disponibilité de grande masse de données plus une durée d'apprentissage considérable. Plusieurs techniques et stratégies ont été introduites en se basant sur des réseaux/modèles pré-entraîné, nous pouvons résumer ces techniques dans les deux suivants :

- **Extraire les caractéristiques avec un réseau de neurones convolutifs** : Les réseaux de neurones profonds sont des architectures hautement configurables avec divers hyperparamètres, les couches initiales capturent des caractéristiques génériques, tandis que les dernières se concentrent davantage sur la tâche spécifique à accomplir (dans notre cas la classification). Cette architecture nous permet d'utiliser la partie du réseau pré-entraîné chargé par l'extraction des caractéristique pour une autre tâche de classification, et de remplacer la dernière couche de classification par un autre classificateur, comme le montre le diagramme suivant :

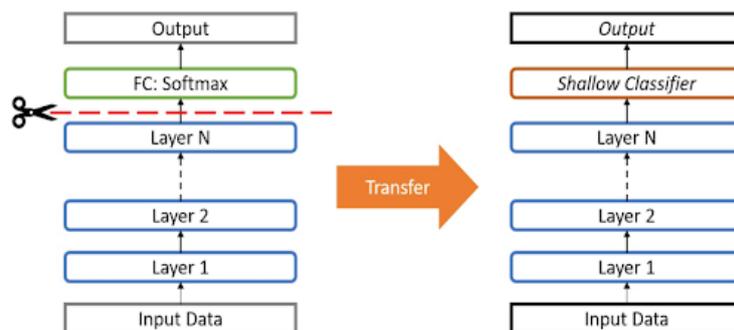


FIGURE 2.14 – Extraire les caractéristiques avec un CNN [?].

- **Utiliser un réseau de neurones convolutifs pré-entraîné** : Il s'agit d'une technique plus complexe, dans laquelle nous utilisons des modèles pré-entraînés tels que : VGG16, ResNet50, etc. Dans cette technique, nous ne contentons pas de remplacer la couche finale (pour la classification), mais également à affiner les poids du réseau pré-entraîné en continuant la rétropropagation. Il est possible d'affiner toutes les couches du réseau convolutif ou de conserver certaines des couches précédentes fixes.
- (c) **Les réseaux de neurones convolutifs pré-entraînés populaires** : Étant donné que les réseaux de neurones convolutifs modernes prennent deux à trois semaines pour l'apprentissage sur plusieurs GPU, plusieurs équipes de recherches partagent leurs résultats sur les réseaux de neurones, ainsi que les paramètres et poids obtenus. Les modèles les plus connus, et qui sont utilisés dans ce mémoire sont :
- **VGG-16** : VGG-16 est constitué de plusieurs couches, dont 13 couches de convolution et 3 fully-connected. Il doit donc apprendre les poids de 16 couches. Il prend en entrée une image en couleurs de taille 224×224 px et la classe dans une des 1000 classes. Il renvoie donc un vecteur de taille 1000, qui contient les probabilités d'appartenance à chacune des classes [54]. L'architecture de VGG-16 est illustrée par les schémas ci-dessous :

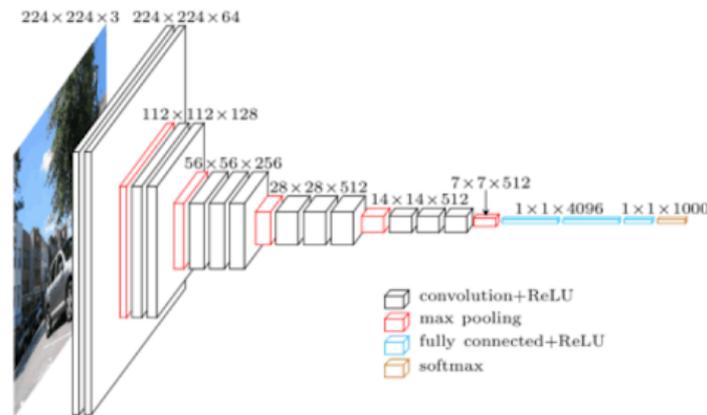


FIGURE 2.15 – l'architecture de VGG-16 [54].

- **ResNet 50** : Contrairement aux architectures de réseau séquentielles traditionnelles telles que VGG, ResNet est plutôt une forme «d'architecture exotique» qui repose sur des modules de micro-architecture (également appelés «architectures réseau dans réseau»). Le terme micro-architecture désigne l'ensemble des «blocs de construction» utilisés pour construire le réseau. Un ensemble de blocs de construction de micro-architecture (avec les couches standard CONV, POOL, etc.) mène à la macro-architecture (c'est-à-dire, le réseau d'extrémité lui-même)[34]. L'architecture de ResNet50 est illustrée par les schémas ci-dessous :



FIGURE 2.16 – l’architecture de ResNet50 [34].

- **MobileNet** : Les réseaux MobileNets sont une classe de réseaux de neurones à convolution conçus par des chercheurs de Google. Ils sont qualifiés de «mobiles d’abord» en ce sens qu’ils ont été conçus de manière à ce qu’ils soient conviviaux en termes de ressources et fonctionnent rapidement, directement sur votre téléphone. La principale différence entre l’architecture MobileNet et un réseau de neurones convolutifs «traditionnel» est qu’au lieu d’une couche unique de convolution $3 \times 3 \times 3$ suivie de de ReLU, les MobileNets divisent la convolution en une convolution de profondeur 3×3 et une convection ponctuelle de 1×1 [36]. L’architecture de MobileNet est illustrée par les schémas ci-dessous :

Type / Stride	Filter Shape	Input Size	
Conv / s2	$3 \times 3 \times 3 \times 32$	$224 \times 224 \times 3$	
Conv dw / s1	$3 \times 3 \times 32$ dw	$112 \times 112 \times 32$	
Conv / s1	$1 \times 1 \times 32 \times 64$	$112 \times 112 \times 32$	
Conv dw / s2	$3 \times 3 \times 64$ dw	$112 \times 112 \times 64$	
Conv / s1	$1 \times 1 \times 64 \times 128$	$56 \times 56 \times 64$	
Conv dw / s1	$3 \times 3 \times 128$ dw	$56 \times 56 \times 128$	
Conv / s1	$1 \times 1 \times 128 \times 128$	$56 \times 56 \times 128$	
Conv dw / s2	$3 \times 3 \times 128$ dw	$56 \times 56 \times 128$	
Conv / s1	$1 \times 1 \times 128 \times 256$	$28 \times 28 \times 128$	
Conv dw / s1	$3 \times 3 \times 256$ dw	$28 \times 28 \times 256$	
Conv / s1	$1 \times 1 \times 256 \times 256$	$28 \times 28 \times 256$	
Conv dw / s2	$3 \times 3 \times 256$ dw	$28 \times 28 \times 256$	
Conv / s1	$1 \times 1 \times 256 \times 512$	$14 \times 14 \times 256$	
5×	Conv dw / s1	$3 \times 3 \times 512$ dw	$14 \times 14 \times 512$
	Conv / s1	$1 \times 1 \times 512 \times 512$	$14 \times 14 \times 512$
Conv dw / s2	$3 \times 3 \times 512$ dw	$14 \times 14 \times 512$	
Conv / s1	$1 \times 1 \times 512 \times 1024$	$7 \times 7 \times 512$	
Conv dw / s2	$3 \times 3 \times 1024$ dw	$7 \times 7 \times 1024$	
Conv / s1	$1 \times 1 \times 1024 \times 1024$	$7 \times 7 \times 1024$	
Avg Pool / s1	Pool 7×7	$7 \times 7 \times 1024$	
FC / s1	1024×1000	$1 \times 1 \times 1024$	
Softmax / s1	Classifier	$1 \times 1 \times 1000$	

FIGURE 2.17 – l’architecture de MobileNet [36].

2.3 La validation croisée

La validation croisée fait probablement partie des techniques les plus importantes utilisées par les scientifiques, car il est toujours nécessaire de valider la stabilité du modèle d'apprentissage automatique. La validation croisée est une technique de validation de modèle permettant d'évaluer la manière dont les résultats d'une analyse statistique (modèle) se généraliseront à un ensemble de données indépendant. Elle est principalement utilisée pour estimer la précision avec laquelle un modèle prédictif fonctionnera, dans des contextes où l'objectif est la prédiction.

La validation croisée a pour objectif de définir un ensemble de données permettant de tester le modèle en phase d'apprentissage (c'est-à-dire un ensemble de données de validation) afin de limiter les problèmes de sur-apprentissage, et d'obtenir un aperçu de la manière dont le modèle se généralisera en un ensemble de données indépendant.

- **Les stratégies de la validation croisée [27]** : En règle générale, différentes stratégies de validation existent en fonction du nombre de divisions effectuées dans l'ensemble de données.
 - (a) **testset validation** : nous divisons l'échantillon de taille n en deux sous-échantillons, le premier dit d'apprentissage (communément supérieur à 60% de l'échantillon) et le second dit de test. Le modèle est entraîné sur l'échantillon d'apprentissage et validé sur l'échantillon de test. L'erreur est estimée en calculant un test, une mesure ou un score de performance du modèle sur l'échantillon de test.
 - (b) **k-fold** : nous divisons l'échantillon original en K échantillons, puis nous sélectionnons un des K échantillons comme ensemble de validation et les $K-1$ autres échantillons constitueront l'ensemble d'apprentissage. Nous calculons comme dans la première méthode le score de performance, puis nous répétons l'opération en sélectionnant un autre échantillon de validation parmi les $K-1$ échantillons qui n'ont pas encore été utilisés pour la validation du modèle. L'opération se répète ainsi K fois pour qu'en fin de compte chaque sous-échantillon ait été utilisé exactement une fois comme ensemble de validation. La moyenne est enfin calculée pour estimer l'erreur de prédiction.
 - (c) **leave-one-out** : Il s'agit d'un cas particulier de Kfold lorsque K est égal au nombre d'échantillons de notre jeu de données. Cela signifie qu'il parcourra chaque échantillon de notre ensemble de données à chaque fois en utilisant un objet $k-1$ en tant qu'échantillons d'apprentissage et un seul objet en tant qu'ensemble de test.

Cette méthode peut être utile si nous avons trop peu de données et si le modèle est assez rapide pour se recycler.

2.4 Conclusion

Dans ce chapitre, nous avons exploré les différentes méthodes d'indexation d'images existantes. Nous avons vu les principes généraux de l'indexation d'images, l'architectures de ces systèmes, ainsi que les différentes approches classiques et adaptatives qui permettent de classifier les images. Dans le prochain chapitre, nous exposerons la méthodologie que nous avons suivie dans notre travail, et les architectures des solutions proposées.

Chapitre 3

Système de classification d'images pour la gestion des déchets

3.1 Introduction

Après avoir présenté une vue globale sur l'existant dans le domaine des Content-Based Image Retrieval (CBIR), ainsi que l'apprentissage profond, nous entamons nos contributions étape par étape en détaillant chacune d'elles avec des exemples illustratifs pour mieux visualiser les démarches. Ceci est fait en se basant sur les connaissances acquises de l'étude réalisée dans le chapitre 2.

3.2 Contexte et motivation

Le développement économique durable, l'excellente qualité de vie ainsi que la gestion éclairée des ressources naturelles, tous ces avantages garanties par les villes intelligentes [16], impulsent la plus part des pays tels que les pays d'union Européen de transformer leurs villes en des villes intelligentes [24].

Dans le cadre de la transformation des ville traditionnelle à des villes intelligentes, nous proposons des approches d'automatisation de la gestion des déchets municipaux. Pour commencer, nous avons mené une étude comparative sur les solutions de classification d'images existants pour la gestion des déchets solides [31], [32].

A partir de cette étude, nous avons constaté que les solutions proposées pour résoudre le problème de classification d'images pour la gestion des déchets solides se répartissent en deux types : le premier type englobe les solutions à base de système de recherche classique, le deuxième type représente l'ensemble des solutions basées sur le système de recherche adaptative. Nous avons remarqué aussi que ces solutions ont aboutit à des résultats qui convergent. Cela nous a motivé à établir deux nouveaux systèmes : Un basé sur la recherche classique et l'autre sur la recherche adaptative puis faire une étude comparative entre eux afin d'extraire la meilleur solution.

3.3 Solution à base de système de recherche classique

Dans cette section, nos contributions sont basées sur des approches classiques et d'autres adaptatives, la section actuelle est dédiée à la présentation de la solution basée sur le système de recherche classique. Cette solution comprend trois phases : la première est la phase hors ligne dans laquelle nous pré-traitons nos images, puis nous extrairons les caractéristiques d'images en utilisant l'algorithme GLCM.

La deuxième est la phase en ligne : L'utilisateur choisit l'image (la requête) à classer, le système extrait les caractéristiques de la requête en question en suivant les mêmes démarches de la phase 1.

Enfin nous passons à la dernière phase (phase 3) : Il s'agit de catégoriser l'image par le classificateur K-ppv. Lors de l'implémentation de l'algorithme K-ppv, nous avons testé trois distances de similarité différentes : Euclidien, Cosine et Corrélacion.

L'utilisation de plusieurs mesures de similarité avec l'algorithme K-ppv nous permet d'établir une étude comparative. En se basant sur les résultats obtenus, le meilleur classificateur adapté au problème de classification d'images pour la gestion des déchets est sélectionné.

L'architecture de l'approche proposée, dans ce travail, à base de système de recherche classique est illustrée dans la figure 3.1.

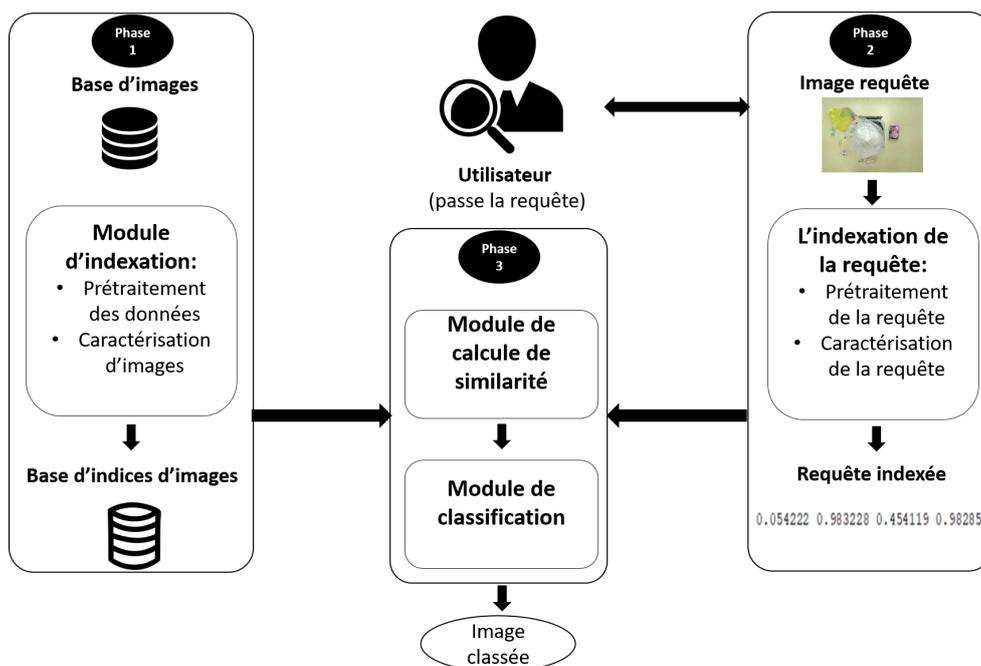


FIGURE 3.1 – L'architecture de l'approche proposée à base de système de recherche classique

Dans ce qui suit, nous allons présenter en détails les différentes phases de l'architecture de l'approche proposée à base de système de recherche classique.

3.3.1 Phase 1 : Module d'indexation d'images

Cette phase se fait en mode hors ligne. Elle est effectuée une seule fois dans le processus. Il s'agit de prendre en entrée notre base d'images qui sont sous format RGB, faire le pré-traitement, puis l'extraction des caractéristiques de chaque image afin d'obtenir une base d'images indexées. La caractéristique visuelle (de bas niveau) représente l'index (où signature) de l'image. Il permet de la récupérer lors de la phase de recherche.

(a) **Étape 1** : Pré-traitement d'images

Cette étape a pour objective d'améliorer la qualité de l'image. Nous présentons, dans ce qui suit, les techniques de pré-traitement les plus courantes :

- **Ajustement du contraste** : Afin d'accroître la différence entre les différentes parties de l'images, nous avons appliqué l'ajustement du contraste, la figure 2.1 illustre l'effet d'ajustement du contraste.
- **Conversion de l'image aux niveaux de gris** : Les pixels des images de notre base d'apprentissage appartiennent à l'espace RGB. Dans la phase de prétraitement et après l'ajutement du constrast, nous allons les convertir aux niveaux de gris pour enfin les soumettre au GLCM.



FIGURE 3.2 – Exemple illustre l'effet d'ajustement du contraste

A la fin de cette étape nous avons obtenus des images ajustées et converties aux niveaux de gris. À partir des images améliorées, dans l'étape de pré-traitement, les caractéristiques visuelles sont extraites

(b) **Étape 2** : Caractérisation d'images :

Il s'agit de représenter chaque image de notre base de données sous forme d'un vecteur de caractéristiques. Ce vecteur est considéré comme la nouvelle représentation d'image qui encode la représentation initiale sous forme de matrice de pixels. Cette représentation nous permet d'effectuer la recherche d'images similaires. La figure 3.3 présente le processus d'extraction des caractéristiques à partir des images.

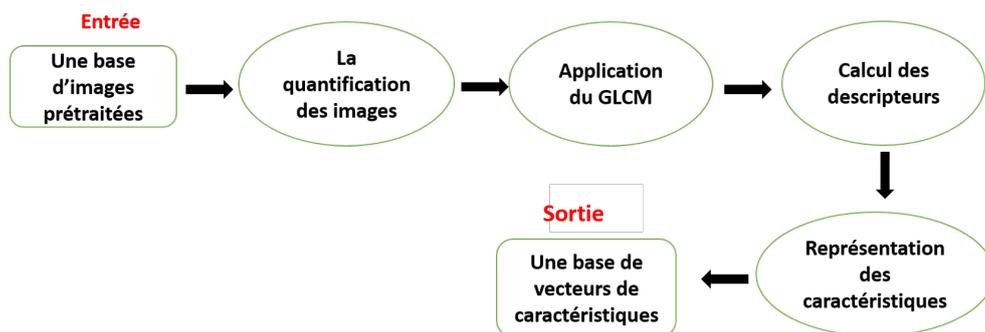


FIGURE 3.3 – Processus d'extraction des caractéristiques à partir des images

Dans ce qui suit, nous détaillons ce processus commençant par la quantification d'images pré-traitées prises en entrée :

— **La quantification d'images** : Elle consiste à réduire les niveaux de gris, cela permet de minimiser les calculs ainsi que la réduction d'espace mémoire occupé ; il s'agit d'effectuer une conversion de l'image d'origine vers une image avec n niveaux de gris, nous avons utilisé une fois $n = 8$ niveaux, une autre fois $n = 16$ niveaux et au final nous avons comparé entre les résultats obtenus pour extraire la meilleur quantification entre eux.

— Pour $n = 8$ nous avons effectué les remplacements suivantes :

- niveau 0 $\in [0, 32[$.
- niveau 1 $\in [32, 64[$.
- niveau 2 $\in [64, 96[$.
- niveau 3 $\in [96, 128[$.
- niveau 4 $\in [128, 160[$.
- niveau 5 $\in [160, 192[$.
- niveau 6 $\in [192, 224[$.
- niveau 7 $\in [224, 256[$.

— Et pour $n = 16$ nous avons effectué ces remplacements :

- niveau 0 $\in [0, 16[$.
- niveau 1 $\in [16, 32[$.
- niveau 2 $\in [32, 48[$.
- niveau 3 $\in [48, 64[$.
- niveau 4 $\in [64, 80[$.
- niveau 5 $\in [80, 96[$.
- niveau 6 $\in [96, 112[$.
- niveau 7 $\in [112, 128[$.
- niveau 8 $\in [128, 144[$.
- niveau 9 $\in [144, 160[$.
- niveau 10 $\in [160, 176[$.
- niveau 11 $\in [176, 192[$.
- niveau 12 $\in [192, 208[$.
- niveau 13 $\in [208, 224[$.
- niveau 14 $\in [224, 240[$.
- niveau 15 $\in [240, 256[$.

La figure 3.4 démontre un exemple de quantification d'une partie d'une image à base de 8 niveaux de gris.

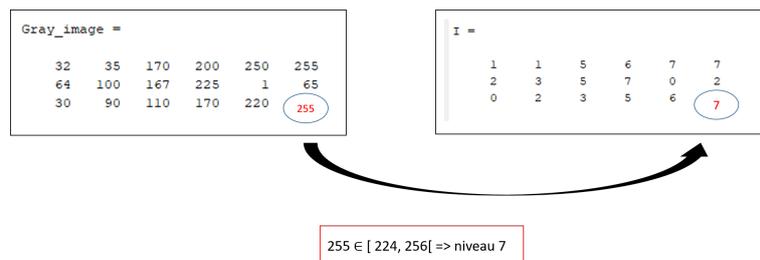


FIGURE 3.4 – Quantification d'une image

— **L'application de l'algorithme GLCM** : Après avoir quantifier les images nous avons construit une matrice carrée de taille $n * n$ où n représente le nombre de niveaux de gris de l'images quantifiée et l'avons rempli selon l'algorithme GLCM déjà cité dans chapitre 2. Nous avons fixé la distance à $d = 1$ et l'angle à $\theta = 0$ pour parcourir toute la matrice et diminuer la perte d'information.

La figure 3.5 illustre une matrice de co-occurrence obtenue à partir de l'application d'algorithme GLCM sur une image de notre base d'images du système.

```

glcm =
      0      11      3      1      1      0      0      0
    15    3447    1372    264    154    41    16    2
      1    1329    17480    1479    337    102    31    16
      0     347    1468    4805    1153    266    119    27
      0     123     286    1228    9295    1477    317    75
      0      38     120     272    1405    74137    3709    248
      0      13      34     100     385    3837    295592    3303
      0       3      12      36      71     275    3274    45448

```

FIGURE 3.5 – Exemple du GLCM appliqué sur une image de la base d’images du système

- **Calcul des descripteurs statistiques :** A partir de la matrice de co-occurrence, nous avons effectué plusieurs statistiques, chaque statistique est considérée comme une caractéristique de l’image; le tableau suivant montre les valeurs obtenues après l’application numérique des formules : (2.1), (2.2), (2.3) et (2.5) dans l’ordre sur la matrice de la figure 3.5 :

Statistique	Valeur
Contraste	0.1038
Corrélation	0.9613
Énergie	0.4152
Uniformité	0.9679

TABLE 3.1 – Application numérique des descripteurs statistiques

- **Représentation des caractéristiques :** Les descripteurs statistiques calculés sont rassemblés dans un seul vecteur. Ce dernier devient l’index de l’image. La figure montre quelques index extrait à partir d’un sous ensemble d’images de notre base d’apprentissage.

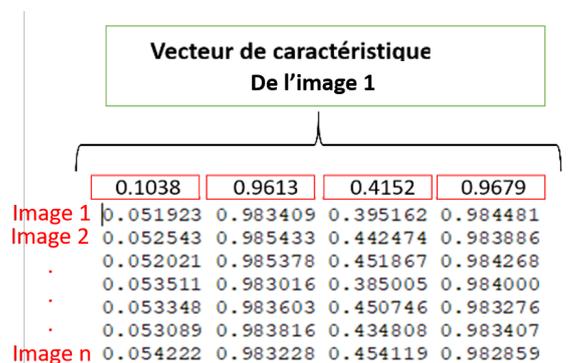


FIGURE 3.6 – Exemple de quelques vecteurs de caractéristiques

3.3.2 Phase2 : L’indexation de la requête

Cette phase suit les mêmes étapes citées dans la phase 1 (section 3.3.1) qui sont : Lors du pré-traitement, le contraste de l’image requête est ajustée puis elle est convertie au

niveaux de gris. Une fois le prétraitement effectué, les caractéristiques visuelles sont extraites à partir de l'image améliorée.

3.3.3 Phase 3 : Module de classification

En se basant sur les descripteurs visuels, l'image est assignée à une de ces classes "Low", "Medium", "Full", "Flow" ou "Overflow". Pour cela, nous appliquons un algorithme de classification supervisée, il s'agit de l'algorithme "**K-plus proches voisins (K-ppv)**". Ce dernier exige à avoir en entrée trois paramètres qui sont :

- (a) La base d'apprentissage : L'ensemble d'images indexées, présentées sous forme de vecteurs de caractéristiques, ainsi que leurs classes à lesquelles elles appartiennent.
- (b) La requête présentée sous le même format que la base d'apprentissage. Autrement dit, la même méthode d'indexation doit être appliquée à l'image requête.
- (c) Le nombre de K plus proches voisins.

Pour chaque image requête, l'algorithme K-ppv donne, en sortie la classe ciblée. Pour cela, le K-ppv passe par les étapes suivantes :

- (a) **Calcul de similarité** : Cette étape consiste à calculer la distance de similitude des images de la base d'apprentissage et de la requête. Ce calcul est basé sur la comparaison des vecteurs de caractéristiques attribué aux images lors de la phase d'indexation. Plusieurs distances de similarité peuvent être utilisées par exemple "cosine", "euclidienne"....etc. Dans ce travail, nous avons exploré trois distances différentes : cosinus, euclidienne et corrélation avec le classificateur K-ppv.

Ci-après, nous présentons un exemple d'application de la mesure de similarité. Il s'agit de calculer la "distance cosinus" entre les index d'une image de la base d'apprentissage et l'index de la requête :

$$\mathbf{I} = [0.094180 \quad 0.984743 \quad 0.799547 \quad 0.988200].$$

$$\mathbf{Requ\^ete} = [0.164960 \quad 0.956753 \quad 0.669804 \quad 0.969925].$$

En l'appliquant de la formule (2.7), nous obtenons le résultat suivant :

$$d_{(I, Requ\^ete)} = 0.0032.$$

- (b) **Le classement des images retournées par l'algorithme K-ppv** : Les images retournées par l'algorithme de classification sont ordonnées selon les distances de similarité. Les images les plus proches en terme de similarité sont classées en premier.
- (c) **L'extraction des k premiers voisins** : A partir de la liste des images triées en se basant sur la distance de similarité, nous extrairons les k premières images qui sont considérées comme les plus proches voisins de la requête.
- (d) **Le calcul de la classe majoritaire** : Chaque image voisine dispose d'une étiquette qui indique sa classe. Nous calculons depuis ces étiquettes le nombre d'apparition de chaque

classe. La classe qui a une occurrence élevée est dite classe majoritaire. L'exemple suivant explique comment déterminer la classe majoritaire.

Étant donnée que nous avons 5 images considérées comme les plus proches voisins ($k = 5$) :

La classe du premier voisin est : **"Flow"**

La classe du seconde voisin est : **"OverFlow"**

La classe du troisième voisin est : **"Flow"**

La classe du quatrième voisin est : **"Flow"**

La classe du cinquième voisin est : **"Full"**

En calculant le nombre d'apparition de chaque classe nous obtenons :

Classe	Nombre d'occurrence
Low	0
Medium	0
Full	1
Flow	3
OverFlow	1

TABLE 3.2 – Nombre d'occurrence des classes

Selon le tableau 3.2 la classe majoritaire pour l'exemple précédent est : **"Flow"**.

- (e) **Assigner une classe à la requête** : A la fin de cette phase, la classe majoritaire est la classe assignée à la requête. Dans l'exemple précédent, la classe de la requête est **"Flow"**.

La figure 3.7 résume le processus de classification.

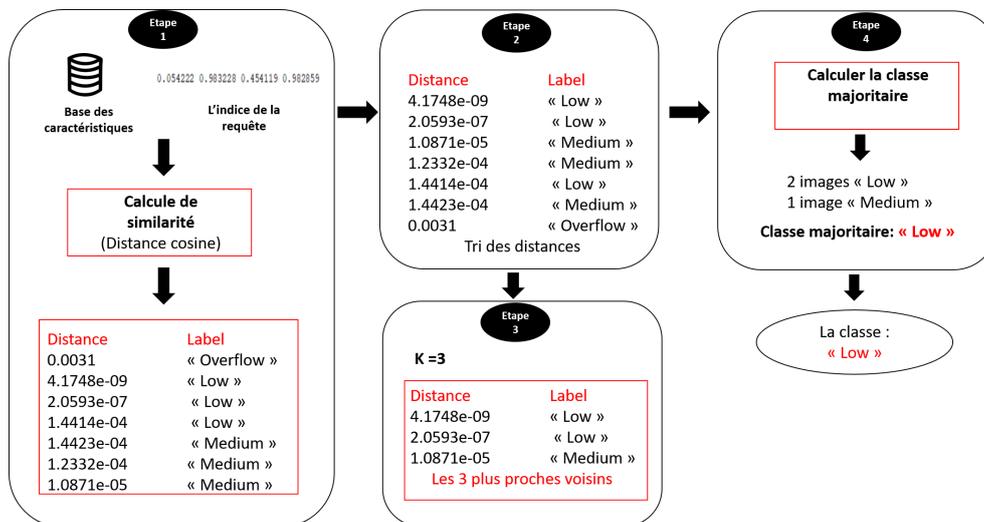


FIGURE 3.7 – Schéma détaillé de la phase de classification

Nous avons présenté, dans cette section, la conception de notre approche de recherche d'images sur les systèmes classique ainsi que ses différents phases et modules. La section suivante est consacrée à la présentation de la conception de nos approches basées sur le système de recherche adaptative.

3.4 Solution à base de système de recherche adaptative

Dans cette section, nous allons montrer la deuxième solution à base des systèmes de recherche adaptative. Ces solutions sont basées, principalement, sur l'apprentissage profond ainsi que la technologie d'apprentissage par transfert. Nous nous intéressons particulièrement à deux principales stratégies : l'extraction des caractéristiques et le fine-tuning d'un réseau pré-entraîné. Afin d'obtenir des résultats fiables et efficaces, nous avons opté pour une étude comparative entre ces deux stratégies, nous avons donc implémenté un système pour chaque stratégie. Dans ce qui suit, nous allons détailler les démarches suivies pour atteindre cet objectif.

3.4.1 Méthodologie et conception du système de classification basé sur le fine-tuning :

Dans ce système, nous allons explorer la technique fine-tuning qui permet d'exploiter un réseau de neurones convolutifs pré-entraîné, et l'adapter à un nouveau cas de classification. Le système est composé de deux phases, chaque phase contient un module. La première phase dite phase hors ligne est consacrée à l'apprentissage du modèle. Cette phase débute par une étape essentielle qui est le pré-traitement de la base d'images. Nous obtenons par la suite une nouvelle base d'images pré-traitées. Ces dernières passent par deux autres étapes :

- (a) La première étape concerne l'adaptation de l'architecture du réseau pré-entraîné au nouveau cas de classification. Cela est effectué en changeant l'architecture du réseau. Dans cette étape, nous explorons trois différentes architectures de réseaux de neurones convolutifs à savoir le VGG-16, MobileNet, ResNet-50.
- (b) La deuxième étape consiste à l'entraînement du réseau de neurones convolutifs sur la base d'apprentissage. Durant cette étape, les paramètres du réseau sont ajustés au nouveau problème. Enfin, après l'évaluation des performances, le modèle entraîné est déployé.

A l'arrivée d'une nouvelle requête (image), la deuxième phase du système se déclenche pour faire le traitement de la requête suivant la même étape de processus de pré-traitement appliqué sur la base d'apprentissage. Par la suite, l'image doit être classifiée. Pour cela, le module de classification reçoit en entrée le modèle classifieur conçu auparavant dans la phase hors ligne, ainsi que l'image requête pré-traitée. Enfin, une classe est associée à l'image requête. L'architecture de l'approche proposée, dans ce travail, basée sur la technique de fine-tuning est illustrée dans la figure 3.8.

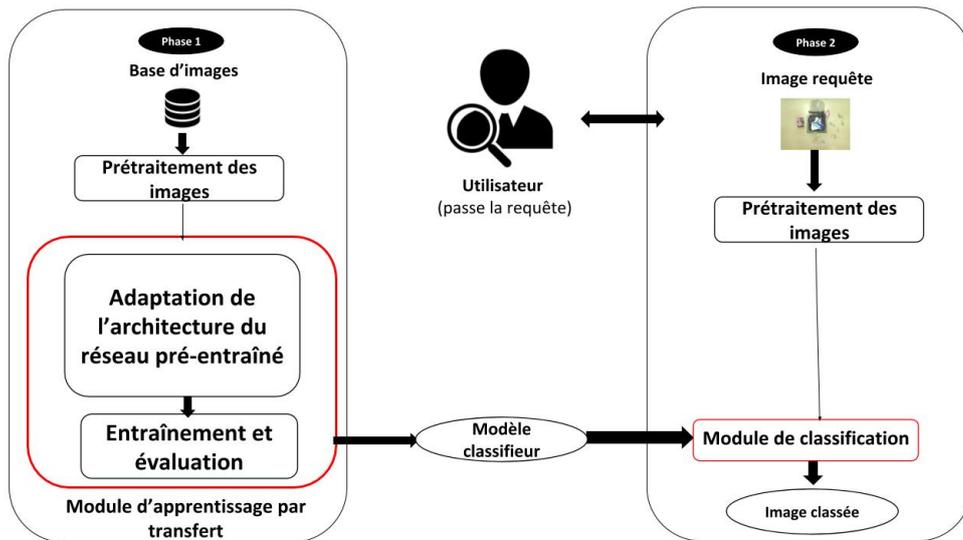


FIGURE 3.8 – L’architecture de l’approche proposée basée sur la technique de fine-tuning.

Dans ce qui suit, nous allons présenter en détails les différentes phases de l’architecture de l’approche proposée à base de fine-tuning.

- (a) **Phase 1 : Hors ligne** Cette phase s’effectue une seule fois dans le processus. Il s’agit de pré-traiter la base d’images, faire l’adaptation de l’architecture du réseau-pré-entraîné, puis le faire passer par l’entraînement et l’évaluation.

— **Étape 1 : Pré-traitement**

La première étape de cette phase, consiste à réaliser un pré-traitement de la base d’images. Les images sont sous format RGB, nous effectuons une conversion au niveaux de gris. Cette étape a pour but de préparer les images pour l’étape d’apprentissage. La figure 3.9 représente un exemple d’une image de la base d’apprentissage sous format RGB, ainsi que le résultat après la conversion au niveaux de gris.

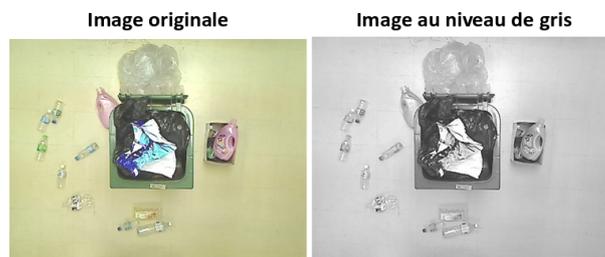


FIGURE 3.9 – Exemple d’une image avant et après le pré-traitement

— **Étape 2 : Module d’apprentissage**

Dans cette étape, nous avons une base d’images pré-traitées, cette dernière est l’entrée du module d’apprentissage, en sortie nous aurons un modèle classifieur. Le module d’apprentissage est composé de deux sous étapes, nous allons en détailler chacune.

a) **Sous étape 1 : Adaptation de l'architecture du réseau pré-entraîné**

Nous avons utilisé des modèles CNN, déjà entraînés sur un grand ensemble de données d'image comme ImageNet (contient environ 1,4 millions d'images étiquetées et 1000 catégories différentes), cela dit que la sortie (la dernière couche) est composée de 1000 noeuds. Dans notre problème de classification, le nombre de classes est 5, représentant les niveaux de quantité des déchets.

La technique implémentée consiste à éliminer la dernière couche du réseau CNN, et la remplacer par une autre avec cinq noeuds (cinq classes).

A la fin de cette étape, nous avons un modèle pré-entraîné adapté à notre problème de classification, nous allons donc l'entraîner et l'évaluer pour enfin obtenir un modèle classifieur.

b) **Sous étape 2 : Entraînement et évaluation**

Cette étape a pour objectif de faire l'apprentissage et la classification du modèle construit dans l'étape précédente. Avant de passer à l'apprentissage, nous avons divisé notre base d'images en deux sous ensembles, un pour l'apprentissage (75%), l'autre pour le test (25%). L'apprentissage se fait à travers les images qui constituent la grande partie de la base, et comme il s'agit d'une méthode supervisée, nous allons passer au modèle les images ainsi que leurs classes.

Pour l'évaluation, le modèle reçoit en entrée les images de test, et il donne en sortie un vecteur de probabilités d'appartenance de l'image à chaque classe.

(b) **Phase 2 : En ligne**

Cette phase est consacrée aux opérations traitant l'arrivée d'une nouvelle image requête. La phase est composée d'une étape de pré-traitement et un module de classification.

— **Étape 1 : Pré-traitement de l'image requête**

L'image requête arrive dans le système sous le format RGB, elle passe par l'étape du pré-traitement qui se fait de la même méthode citée dans la phase 1. En sortie, l'image requête est convertie aux niveaux de gris.

— **Étape 2 : Module de classification**

Cette étape est dédiée à l'attribution de l'image requête à sa classe. Pour cela, nous avons notre modèle classifieur, qui prend en entrée l'image requête pré-traitée, en sortie nous avons un vecteur de probabilités d'appartenance de l'image à chaque classe. Enfin, la classe ayant la plus grande probabilité est attribuée à l'image requête. La figure 3.10 présente le processus de classification.

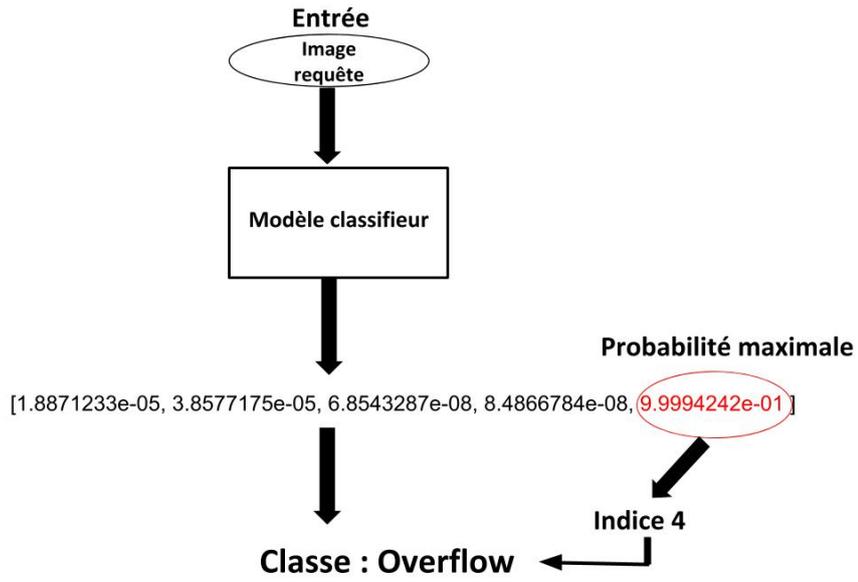


FIGURE 3.10 – Processus de la classification d'une requête.

3.4.2 Méthodologie et conception du système à base d'apprentissage profond pour l'extraction des caractéristiques

Dans ce système, nous avons investigué une autre technique de l'apprentissage par transfert. Cette dernière consiste à utiliser les premières couches d'un modèle pré-entraîné pour faire l'extraction des caractéristiques des images. Autrement dit, nous utilisons la première partie d'un réseau de neurones profond pour représenter l'image. Cette représentation, constituée d'un vecteur de descripteurs de bas niveau, sera l'entrée d'un classificateur supervisé. Le système conçu, exploite trois architectures profondes : VGG-16, ResNet-50, MobileNet pour l'extraction des caractéristiques, et SVM pour la classification.

Le déroulement du système se passe en deux phases principales, une première phase dédiée à la construction et l'apprentissage du modèle de classification. La deuxième phase est consacrée à la classification d'une image requête. La figure 3.11 illustre l'architecture de l'approche proposée. Dans ce qui suit les différents modules de cette architecture sont détaillés.

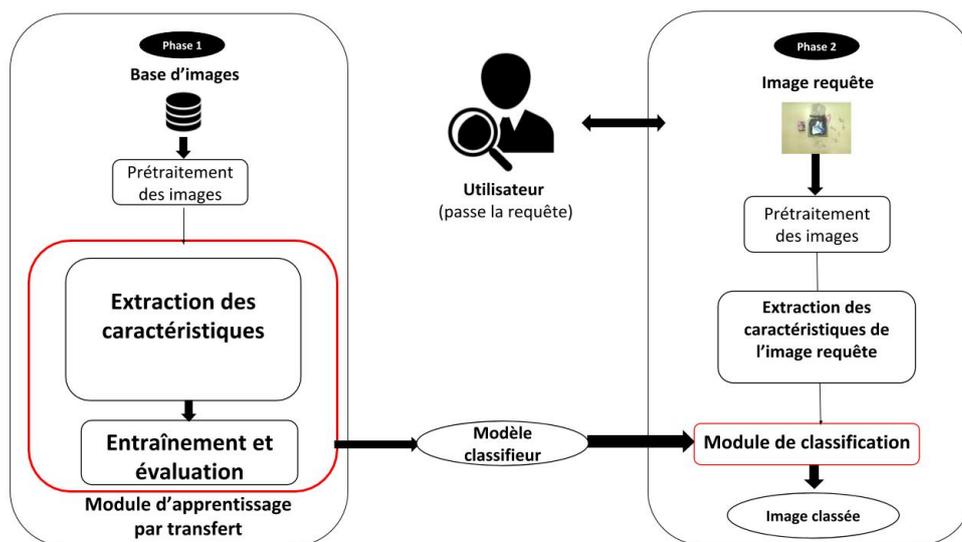


FIGURE 3.11 – L'architecture de l'approche proposée à base d'apprentissage profond pour l'extraction des caractéristiques.

(a) Phase 1 : Hors ligne

— Étape 1 : Pré-traitement

Les images de la base d'apprentissage sont converties du RGB aux niveaux de gris, suivant le processus cité dans la section du pré-traitement du système précédant 4.23.

— Étape 2 : Module d'apprentissage par transfert

Cette étape est chargée de la création et l'évaluation du modèle classifieur, elle est constituée de deux sous étapes :

a) **Sous étape 1 : Extraction des caractéristiques**

L'architecture typique d'un réseau de neurones profond -CNN se divise en deux parties, la première est celle chargée d'extraction de caractéristiques, elle est composée de plusieurs couches de convolutions. La deuxième est celle de classification. Nous nous intéressons dans ce système à la partie d'extraction des caractéristiques. Nous avons implementé cette architecture avec trois variantes de modèles de réseaux de neurones : VGG-16, MobileNet, ResNet-50. Ces modèles sont employés pour l'extraction des caractéristiques des images pré-traitées. Pour chaque image, nous obtenons, une vecteurs de caractéristiques. Dans la nouvelle base d'apprentissage, chaque image est représentée par sa signature de bas niveau qui est considérée comme l'entrée du classificateur SVM.

b) **Sous étape 2 : Entraînement et évaluation**

Dans cette sous-étape, la base d'apprentissage sera divisé en deux partitions : apprentissage et test. La première partition est constituée 75% des images de la base.Elle sert à faire entraîner le modèle de classification. Tandis que la deuxième partition, dédiée à l'évaluation des performances du classificateur, est constituée de 25%. Enfin, le classificateur SVM est prêt pour être déployé.

(b) **Phase 2 : En ligne**

- **Étape 1 : Pré-traitement** Dans cette étape, Le même traitement indiqué dans l'étape hors ligne est appliqué à l'image requête.
- **Étape 2 : Extraction de caractéristiques de l'image requête** Cette étape a pour but d'extraire les caractéristiques de l'image requête pré-traitée, pour cela, nous avons utilisé un modèles pré-entraîné, étant donné le VGG-16. Ce dernier reçoit en entrée l'image aux niveaux de gris, et en sortie il donne le vecteur de caractéristiques représentant l'image au bas niveau.
- **Étape 3 : Module de classification**
Après avoir extraire les caractéristiques de l'image requête, elle passe au classifieur SVM, pour avoir sa classe correspondante.
La figure 3.12 montre les étapes suivit de l'arrivée de la requête jusqu'à l'attribution de classe.

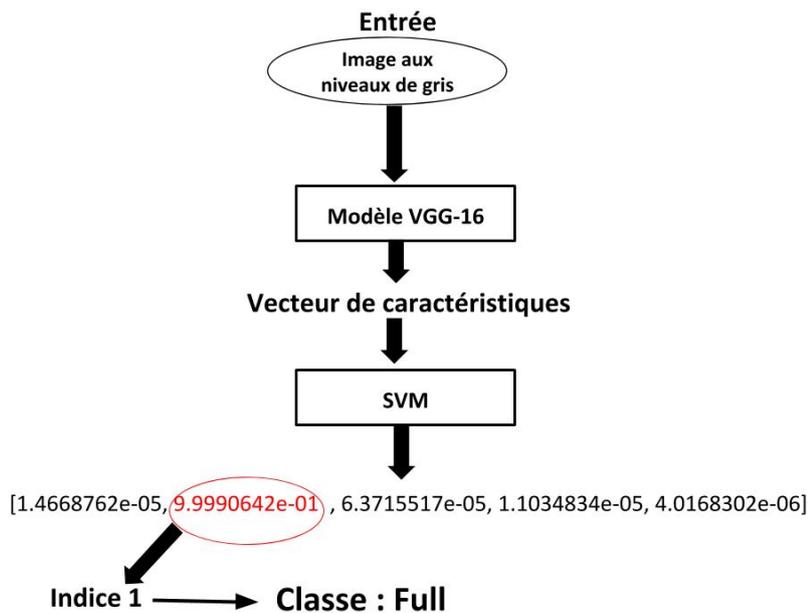


FIGURE 3.12 – Processus de la classification d’une requête.

3.5 Conclusion :

Au cours de ce chapitre, nous venons de présenter les architectures des solutions proposées pour la classification d’images dans un système de gestion de déchets. Ces solutions sont basées sur deux concepts différents, la recherche d’image classique et l’apprentissage par transfert pour la classification d’images. Les solutions proposées seront évalués via des métriques d’évaluation afin de déterminer leurs degré d’efficacité et fiabilité. Les résultats des tests ainsi que la base d’images utilisés seront présentés dans le chapitre suivant.

Chapitre 4

Résultats expérimentaux et évaluation

4.1 Introduction

Dans ce chapitre, nous montrons les expériences et les résultats des tests réalisés afin d'estimer la qualité de notre système et sa performance comparée aux autres systèmes qui existent, ainsi nous discutons les résultats obtenus.

4.2 Démarche expérimentale

4.2.1 Ressources matérielles et logicielles utilisées lors du développement

Nous avons utilisé le matériel et les logicielles suivants :

- (a) **Le matériel** : Nous avons travaillé sur un ordinateur personnel, la figure 4.1 présente sa configuration.

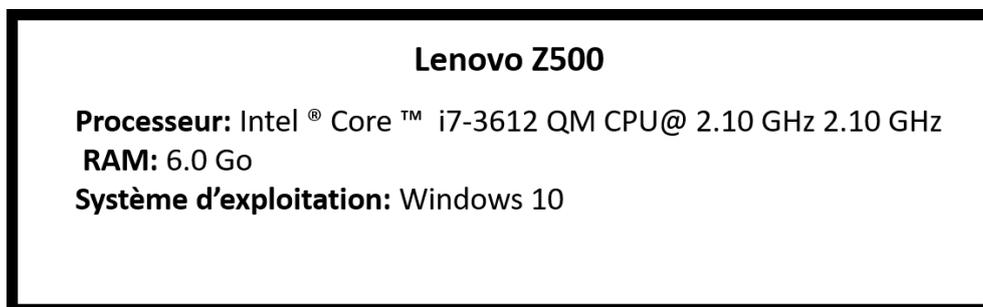


FIGURE 4.1 – Références de l'ordinateur personnel utilisé

- (b) **Présentation du logiciel MATLAB** : MATLAB est le logiciel le plus accessible et le plus productif conçu pour les ingénieurs et les scientifiques. Pour des tâches d'analyse de données, de développement d'algorithmes et de création de modèles, il nous fournit un

environnement qui encourage l'exploration et la découverte. MATLAB associe un langage haut niveau à un environnement bureau élaboré pour les flux de travail itératif scientifique et technique.

Les principales fonctionnalités de MATLAB sont :

- Langage de haut niveau pour le calcul scientifique et technique ;
- Environnement bureau pensé pour l'exploration itérative, la conception et la résolution de problèmes.
- Graphiques destinés à la visualisation de données et outils conçus pour créer des tracés personnalisés.
- Applications dédiées à l'ajustement de courbes, la classification de données, l'analyse de signaux et bien d'autres tâches spécialisées.
- Boîtes à outils additionnelles conçues pour répondre à de nombreux besoins spécifiques aux ingénieurs et aux scientifiques.
- Outils permettant la création d'applications avec interface utilisateur personnalisée.
- Interfaces vers C/C++, Java, .NET, Python, SQL, Hadoop et Microsoft Excel.
- Options de déploiement libre de droits permettant de partager des programmes MATLAB avec les utilisateurs finaux[7].

Nous avons utilisé la version MATLAB R2016a de MATLAB.

(c) **Présentation du langage Python :**

Python est un langage de programmation interprété, multi-paradigme et multiplateformes. Il favorise la programmation impérative structurée, fonctionnelle et orientée objet. Il est doté d'un typage dynamique fort, d'une gestion automatique de la mémoire par ramasse-miettes et d'un système de gestion d'exceptions.

Python est un langage qui peut s'utiliser dans de nombreux contextes et s'adapter à tout type d'utilisation grâce à des bibliothèques spécialisées, pour le développement de notre système de classification d'images à base de système de recherche adaptatif nous avons utilisé la version 3.6.7 du python.

- (d) **La bibliothèque OpenCV :** OpenCV (Open Source Computer Vision Library) est une bibliothèque de logiciels open source de vision informatique et d'apprentissage automatique. Elle a été conçue pour fournir une infrastructure commune aux applications de vision par ordinateur et pour accélérer l'utilisation de la perception de la machine dans les produits commerciaux. OpenCV étant un produit sous licence BSD, il est facile pour les entreprises d'utiliser et de modifier le code.

La bibliothèque contient plus de 2500 algorithmes optimisés, qui incluent un ensemble complet d'algorithmes de vision par ordinateur et d'apprentissage automatique. Ces algorithmes peuvent être utilisés pour détecter et reconnaître des visages, identifier des

objets, classer les actions humaines dans des vidéos, suivre les mouvements d'une caméra, suivre des objets en mouvement, extraire des modèles d'objets 3D, produire des nuages de points 3D à partir de caméras stéréo, assembler des images pour obtenir une haute résolution. l'image d'une scène entière, trouver des images similaires dans une base de données d'images, supprimer les yeux rouges des images prises au flash, suivre les mouvements des yeux, reconnaître les paysages et établir des repères pour les superposer à la réalité augmentée, etc. OpenCV compte plus de 47 000 utilisateurs et le nombre de téléchargements estimé à plus de 18 millions. La bibliothèque est largement utilisée dans les entreprises, les groupes de recherche et les organismes gouvernementaux [5]. Nous avons utilisé la version 4.0.0 d'OpenCV.

(e) **Keras Api** : Keras est une API de réseaux de neurones de haut niveau, écrite en Python et capable de s'exécuter sur TensorFlow, CNTK ou Theano. Elle a été développée pour permettre une expérimentation rapide. Pouvoir faire de la recherche de qualité est essentiel pour pouvoir passer de l'idée au résultat le plus rapidement possible ; Parmi les avantages de l'utilisation de Keras [6] :

- Un prototypage simple et rapide (convivialité, modularité et extensibilité).
- Prendre en charge les réseaux convolutionnels et les réseaux récurrents, ainsi que les combinaisons des deux.
- Fonctionne de manière transparente sur le processeur et le processeur graphique.

4.2.2 Outils développés

Nous avons développé deux applications de classification d'images pour la gestion des déchets municipaux. Dans ce qui suit nous allons présenter l'interface graphique de chaque système proposés.

(a) **Représentation de l'interface graphique du système basé sur la recherche classique** : La figure 4.2 représente l'interface graphique de ce système. Tout d'abord l'utilisateur choisit la distance de similarité, le nombre des K plus proches voisins, ainsi que l'image à classer en cliquant sur le bouton "Load".

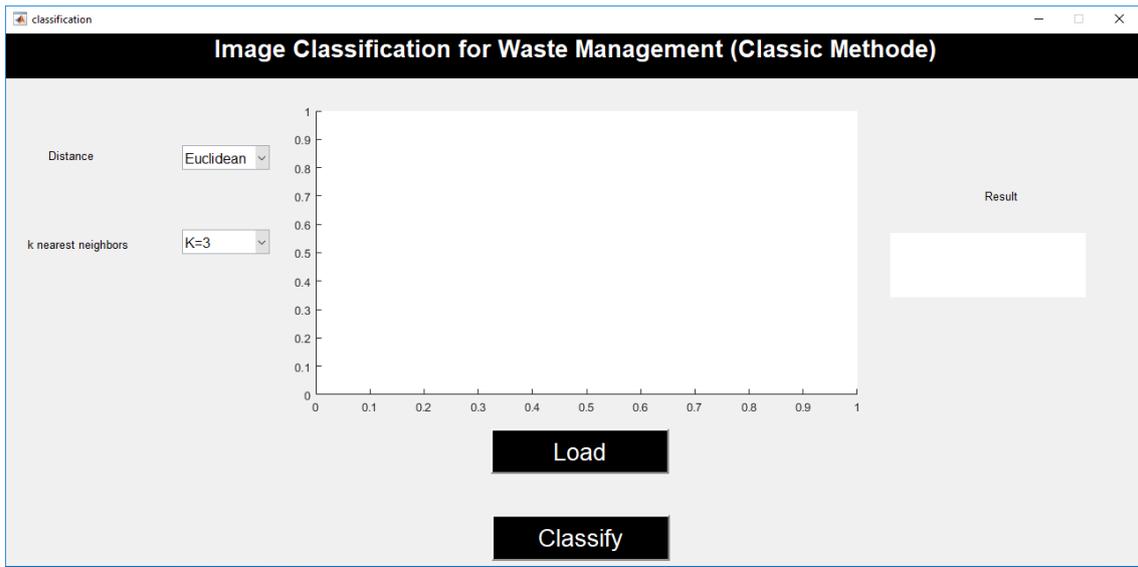


FIGURE 4.2 – Interface d'accueil du système classique

L'image en question s'affiche dans l'axe de l'interface. Puis l'utilisateur clique sur le bouton "Classify" pour que le système affiche la classe de l'image requête dans la zone "Result" et les 6 images similaires à la requête en question dans une nouvelle fenêtre. La figure 4.3 représente l'affichage de résultat de la classification ainsi la figure 4.4 représente la fenêtre présentant les images similaires.

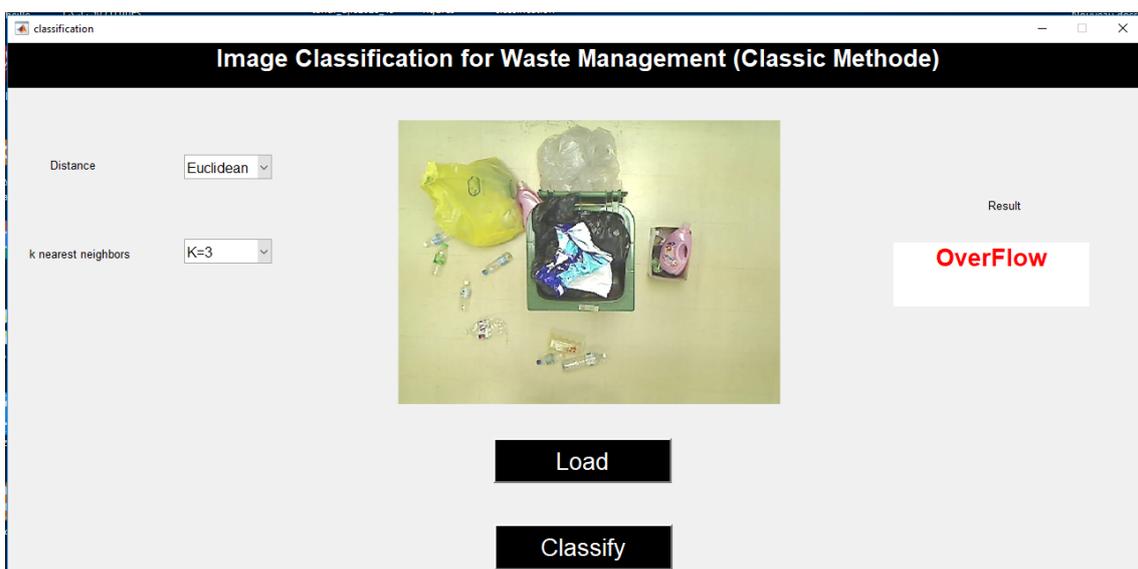


FIGURE 4.3 – Classification de l'image requête

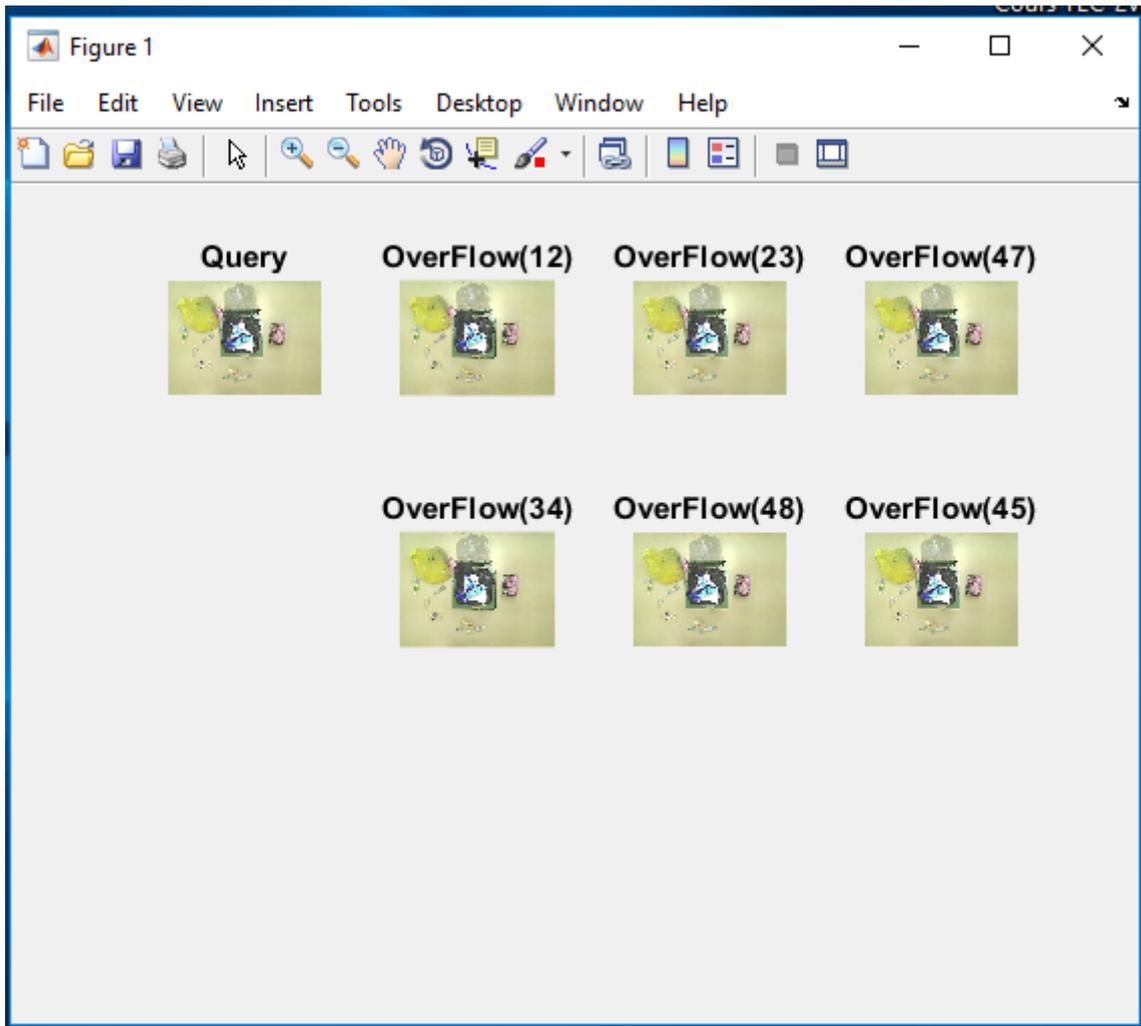


FIGURE 4.4 – Interface des images similaires à l’image requête

(b) **Représentation de l’interface graphique du système basé sur la recherche adaptative**

Dans cette partie, nous présentons l’interface graphique du système adaptative. La figure représente l’interface d’accueil, cette dernière est composée de quatre étapes principales :

- a) L’étape 1 : L’utilisateur choisit une image pour la classifier.
- b) L’étape 2 : L’utilisateur choisit un modèle classificateur, ResNet ou SVM .
- c) L’étape 3 : Après le choix du classificateur, la classe de l’image requête s’affiche dans le champ numéro 3. La figure 4.6 montre un exemple de classification d’une image de classe ‘Medium’.
- d) L’étape 4 : Cette étape consiste à afficher les images similaires à l’image requête, pour cela, l’utilisateur clique sur le bouton ‘display’. Une nouvelle fenêtre s’affiche, comme le montre la figure 4.7

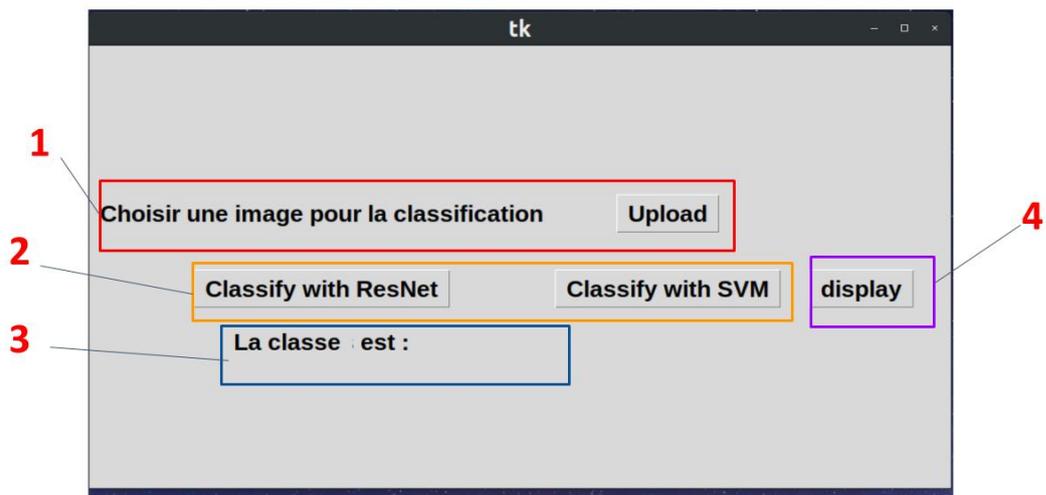


FIGURE 4.5 – Interface d'accueil du système adaptatif

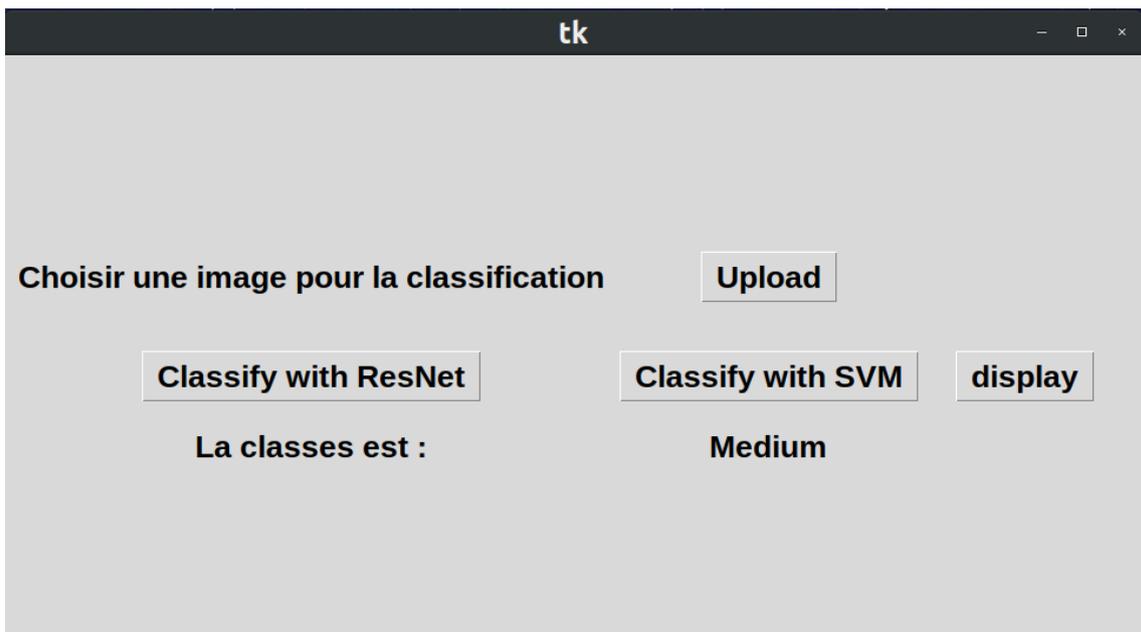


FIGURE 4.6 – Interface d'affichage de la classe de l'image requête

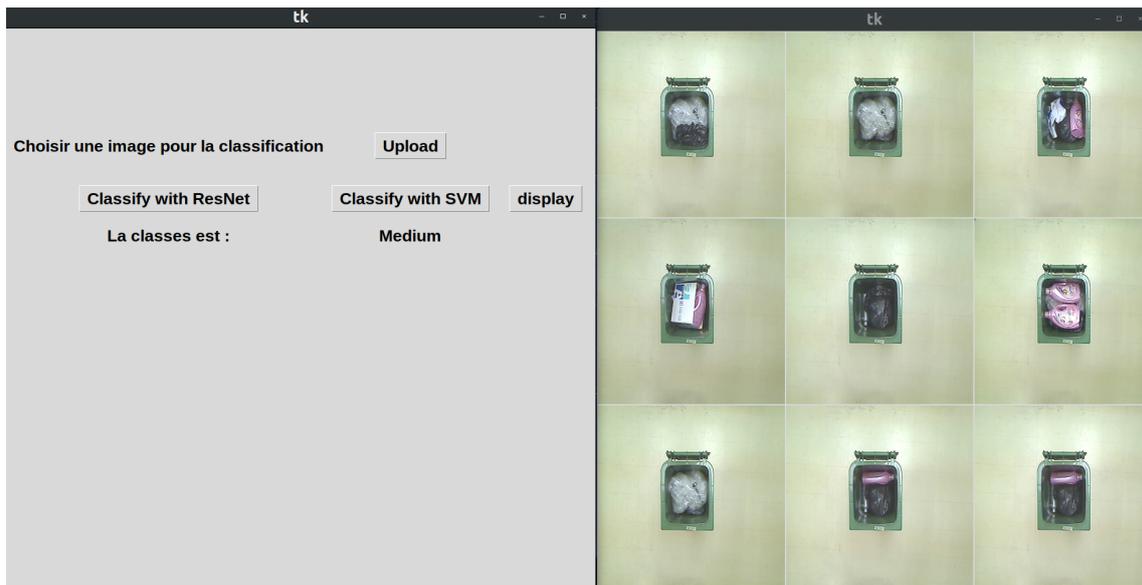


FIGURE 4.7 – Interface des images similaires à l’image requête

4.3 Présentation de la base de données et métriques d’évaluation

4.3.1 Présentation de la base de données

Notre base de données est un ensemble d’images représentant la corbeille à 5 niveaux : "Low", "Medium", "Full", "Flow" et "Overflow", chaque niveau est représenté par 50 images, le total est 250 images prises à l’aide d’une caméra moins coûteuse, de résolution 600*800 Pixel avec 24 milliards de couleurs. Ces images sont prises sous différentes illuminations externes incontrôlées afin d’améliorer la qualité de la collection de base de données. La figure 4.8 illustre des exemples d’images de différentes classes.

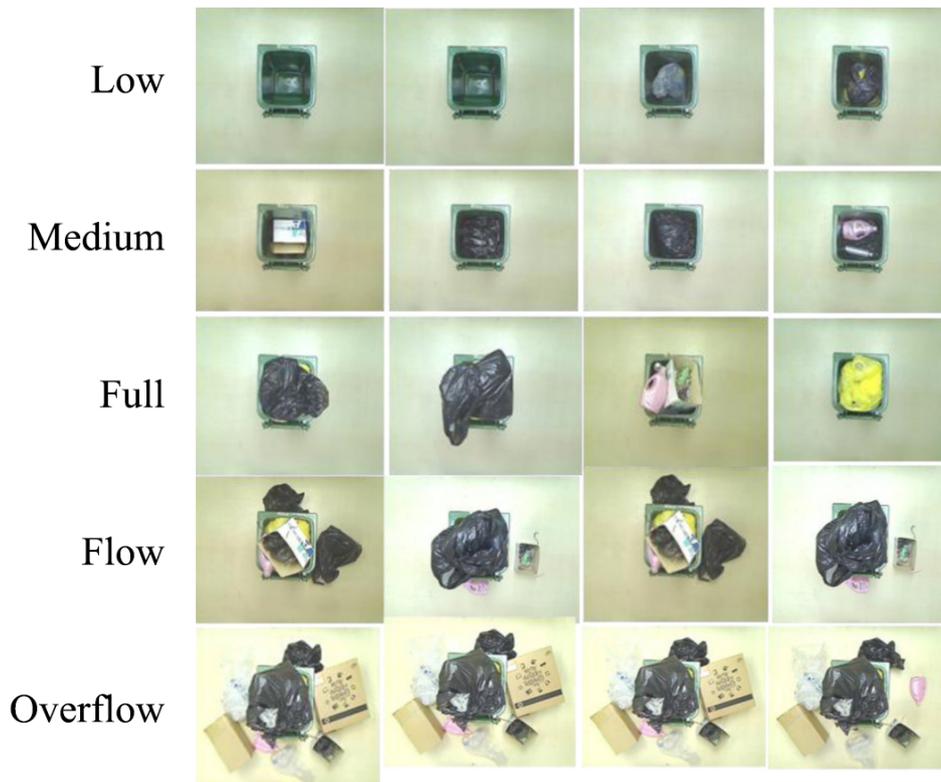


FIGURE 4.8 – Exemple d’images de la base de données des différentes classes.

4.3.2 Métriques d’évaluation

Dans un test de classification, nous classons les éléments selon deux classes : positive ou négative, le résultat peut être correcte ou incorrecte. Pour cela, nous distinguons quatre combinaisons de résultats possibles :

- **Vrais positifs « TP» (True Positive)** : Un résultat est dit vrai positif lorsqu’un élément positif est correctement classé.
- **Faux positifs « FP» (False Positive)** : Un résultat est dit faux positif lorsqu’un élément est classé positif, alors qu’il ne l’est pas.
- **Vrais négatifs « TN» (True Negative)** : Un résultat est dit vrai négatif lorsqu’un élément négatif est correctement classé.
- **Faux négatifs « FN» (False Negative)** : Un résultat est dit faux positif lorsqu’un élément est classé négatif, alors qu’il est positif.

Nous présentons dans la suite les mesures les plus utilisées dans le domaine de RI :

- (a) **La précision** : La proportion de prédictions correctes parmi les points que l’on a prédit positifs

$$Precision = \frac{TruePositive}{TruePositive + FalsePositive} \quad (4.1)$$

(b) **Le rappel : (sensibilité)** C'est le taux de vrais positifs, c'est à dire la proportion de positifs que l'on a correctement identifiés.

$$Rappel = \frac{TruePositive}{TruePositive + FalseNegative} \quad (4.2)$$

(c) **F1-score** : Pour évaluer un compromis entre rappel et précision, on peut calculer la "F-mesure", qui est leur moyenne harmonique.

$$F1 = 2 \times \frac{Precision \times Rappel}{Precision + Rappel} \quad (4.3)$$

(d) **Matrice de confusion :**

Chaque colonne de la matrice représente le nombre d'occurrences d'une classe estimée, tandis que chaque ligne représente le nombre d'occurrences d'une classe réelle (ou de référence). Un des intérêts de la matrice de confusion est qu'elle montre rapidement si un système de classification parvient à classifier correctement [3].

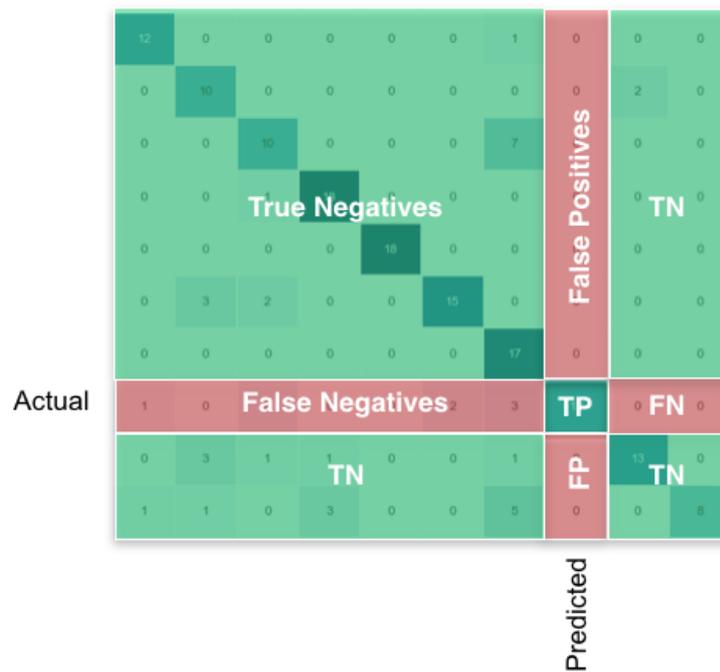


FIGURE 4.9 – Matrice de confusion [3].

4.4 Résultats et discussion

Dans ce qui suit, nous présentons les résultats obtenus dans plusieurs perspectives :

- (a) Les résultats des deux systèmes proposés et de leurs variantes.
- (b) Appliquer une validation croisée pour les deux systèmes, tout en variant le nombre de partitions.
- (c) Tester l'algorithme K-ppv avec plusieurs k possible.
- (d) Tester plusieurs niveaux de gris.

4.4.1 Résultats du système de recherche classique

Dans cette partie, nous présentons nos tests effectués sur la base de test, sans et avec la validation croisée. Nous avons testé l'algorithme K-ppv (en variant la valeur de K) avec les différentes distances implémentées(Cosine, Euclidean et Corrélation). Nous explorons aussi l'effet de variation de nombre des niveaux de gris.

(a) Résultats de l'algorithme K-ppv avec la distance Euclidien

— Sans la validation croisée

Pour la quantification d'images à 8 niveaux de gris :

Plus Proches Voisins	Précision	Rappel	F1-score
$k = 3$	98.57%	98.46%	98.51%
$k = 5$	95.36%	95.12%	95.24%

TABLE 4.1 – Résultats du K-ppv avec 8 niveaux de gris et la distance Euclidien

Pour la quantification d'images à 16 niveaux de gris :

Plus Proches Voisins	Précision	Rappel	F1-score
$k = 3$	97.14%	96.92%	97.03%
$k = 5$	92.61%	91.92%	92.26%

TABLE 4.2 – Résultats du K-ppv avec 16 niveaux de gris et la distance Euclidien

Nous avons constaté que la combinaison d'application du K-ppv où $k=3$ et la distance Euclidien, ainsi que d'utiliser 8 niveaux de gris pour la quantification donne le meilleur résultat avec une précision de 97.14%, rappel de 96.92% et F1-score à 97.03%. La figure 4.10 illustre la matrice de confusion correspondante à ce résultat.

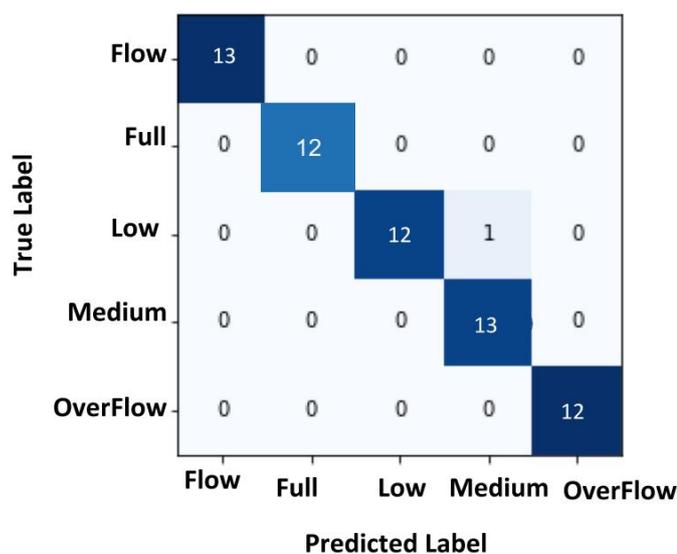


FIGURE 4.10 – Matrice de confusion du 3 plus proches voisins avec la distance Euclidien

Nous avons utilisé la validation croisée par la méthode K-folds afin de mesurer la performance de notre classificateur en variant le partitionnement, une fois le nombre de partitionnement avec deux valeurs 7 et 14. Les 4 tableaux ci-dessous représentent les résultats obtenus.

— Avec la validation croisée

Pour la quantification d’images à 8 niveaux de gris et pour 7 partitions :

Plus Proches Voisins	Précision	Rappel	F1-score
$k = 3$	96.70%	95.51%	96.10%
$k = 5$	92.34%	89.74%	91.02%

TABLE 4.3 – Résultats du K-ppv avec 8 niveaux de gris et la distance Euclidien (avec 7 partitions)

Pour la quantification d’images à 16 niveaux de griset pour 7 partitions :

Plus Proches Voisins	Précision	Rappel	F1-score
$k = 3$	92.17%	89.43%	90.78%
$k = 5$	90.77%	87.24%	88.97%

TABLE 4.4 – Résultats du K-ppv avec 16 niveaux de gris et la distance Euclidien (avec 7 partitions)

Pour la quantification d’images à 8 niveaux de gris et pour 14 partitions :

Plus Proches Voisins	Précision	Rappel	F1-score
$k = 3$	99.07%	98.57%	98.82%
$k = 5$	92.98%	91.78%	92.38%

TABLE 4.5 – Résultats du K-ppv avec 8 niveaux de gris et la distance Euclidien (avec 14 partitions)

Pour la quantification d'images à 16 niveaux de gris et pour 14 partitions :

Plus Proches Voisins	Précision	Rappel	F1-score
$k = 3$	97.92%	97.02%	97.47%
$k = 5$	92.82%	90.11%	91.45%

TABLE 4.6 – Résultats du K-ppv avec 16 niveaux de gris et la distance Euclidien

Nous remarquons que la combinaison de $k=3$, 14 partitions et 8 nombres niveaux de gris est a donné le meilleur résultat en terme de précision (99.07%), rappel (98.57%) et F1-score(98.82%). La figure 4.11 représente la matrice de confusion obtenu par cette combinaison.

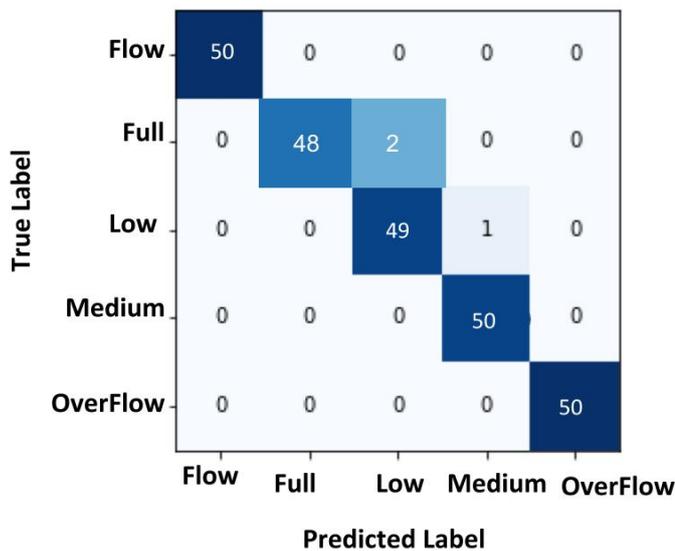


FIGURE 4.11 – Matrice de confusion du 3 plus proches voisins avec la distance Euclidien avec la validation croisée (14 partitions)

(b) Résultats de l'algorithme K-ppv avec la distance Cosine

— Sans la validation croisée

Pour la quantification d'images à 8 niveaux de gris :

Plus Proches Voisins	Précision	Rappel	F1-score
$k = 3$	98.57%	98.46%	98.51%
$k = 5$	95.36%	95.12%	95.24%

TABLE 4.7 – Résultats du K-ppv avec 8 niveaux de gris et la distance Cosine

Pour la quantification d'images à 16 niveaux :

Plus Proches Voisins	Précision	Rappel	F1-score
$k = 3$	98.46%	98.46%	98.46%
$k = 5$	92.61%	91.92%	92.26%

TABLE 4.8 – Résultats du K-ppv avec 8 niveaux de gris et la distance Cosine

— **Avec la validation croisée**

Pour la quantification d'images à 8 niveaux de gris et pour 7 partitions :

Plus Proches Voisins	Précision	Rappel	F1-score
$k = 3$	97.11%	95.91%	96.51%
$k = 5$	92.97%	91.32%	92.14%

TABLE 4.9 – Résultats du K-ppv avec la quantification $n=8$ et la distance Cosine (avec 7 partitions)

Pour la quantification d'images à 16 niveaux de gris et pour 7 partitions :

Plus Proches Voisins	Précision	Rappel	F1-score
$k = 3$	97.92%	97.02%	97.47%
$k = 5$	92.82%	90.11%	91.45%

TABLE 4.10 – Résultats du K-ppv avec 16 niveaux de gris et la distance Cosine (avec 7 partitions)

Pour la quantification d'images à 8 niveaux de gris et pour 14 partitions :

Plus Proches Voisins	Précision	Rappel	F1-score
$k = 3$	99.07%	98.57%	98.82%
$k = 5$	92.98%	91.78%	92.38%

TABLE 4.11 – Résultats du K-ppv avec 8 niveaux de gris et la distance Cosine (avec 14 partitions)

Pour la quantification d'images à 16 niveaux de gris et pour 14 partitions :

Plus Proches Voisins	Précision	Rappel	F1-score
$k = 3$	97.92%	97.02%	97.47%
$k = 5$	92.82%	90.11%	91.45%

TABLE 4.12 – Résultats du K-ppv avec 16 niveaux de gris et la distance Cosine (avec 14 partitions)

Selon les tableaux 4.7, 4.8, 4.9, 4.10, 4.11 et 4.12 nous constatons que les résultats du K-ppv avec la distance Cosine et les résultats du K-ppv avec la distance Euclidien sont similaires.

(c) **Résultats de l’algorithme K-ppv avec la distance Correlation**

— **Sans la validation croisée**

Pour la quantification d’images à 8 niveaux de gris :

Plus Proches Voisins	Précision	Rappel	F1-score
$k = 3$	93.89%	93.46%	93.67%
$k = 5$	93.89%	93.46%	93.67%

TABLE 4.13 – Résultats du K-ppv avec 8 niveaux de gris et la distance Corrélation

Pour la quantification d’images à 16 niveaux de gris :

Plus Proches Voisins	Précision	Rappel	F1-score
$k = 3$	89.15%	88.58%	88.87%
$k = 5$	84.69%	83.84%	84.27%

TABLE 4.14 – Résultats du K-ppv avec 8 niveaux de gris et la distance Corrélation

selon les tableaux 4.13 et 4.14 nous remarquons que la quantification à 8 niveaux de gris et avec $k=3$ donne le meilleur résultat pour le K-ppv et la distance Corrélation, avec une précision de 93.89%, un rappel de 93.46% et un F1-score de 93.67%. La figure 4.12 représente la matrice de confusion obtenue.

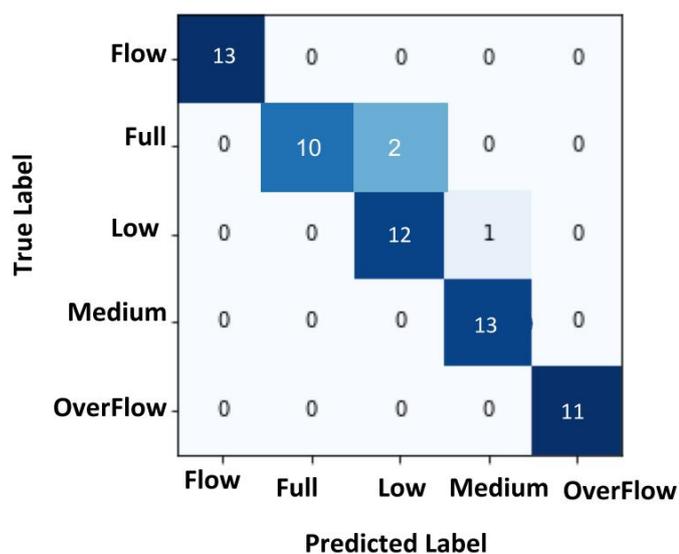


FIGURE 4.12 – Matrice de confusion du 3 plus proches voisins avec la distance Corrélacion sans la validation croisée

— Avec la validation croisée

Pour la quantification d’images à 8 niveaux de gris et pour 7 partitions :

Plus Proches Voisins	Précision	Rappel	F1-score
$k = 3$	95.20%	93.06%	94.12%
$k = 5$	88.78%	87.34%	88.05%

TABLE 4.15 – Résultats du K-ppv avec 8 niveaux de gris et la distance Corrélacion(avec 7 partitions)

Pour la quantification d’images à 16 niveaux de gris et pour 7 partitions :

Plus Proches Voisins	Précision	Rappel	F1-score
$k = 3$	86.23%	85.61%	85.92%
$k = 5$	82.59%	80.40%	81.48%

TABLE 4.16 – Résultats du K-ppv avec 16 niveaux de gris et la distance Corrélacion(avec 7 partitions)

Pour la quantification d’images à 8 niveaux de gris et pour 14 partitions :

Plus Proches Voisins	Précision	Rappel	F1-score
$k = 3$	97.10%	95.71%	96.40%
$k = 5$	92.95%	91.19%	92.06%

TABLE 4.17 – Résultats du K-ppv avec 8 niveaux de gris et la distance Corrélation (avec 14 partitions)

Pour la quantification d'images à 16 niveaux de gris et pour 14 partitions :

Plus Proches Voisins	Précision	Rappel	F1-score
$k = 3$	89.47%	87.26%	88.35%
$k = 5$	87.83%	86.66%	87.24%

TABLE 4.18 – Résultats du K-ppv avec 16 niveaux de gris et la distance Corrélation(avec 14 partitions)

La combinaison de quantification à 8 niveaux et 14 partitions donne le meilleur résultat pour l'application K-ppv($k=3$) avec la distance Corrélation en terme de précision (97.10%), rappel(95.71%) et F1-score(96.40%). La figure 4.13 illustre la matrice de confusion correspondante à cette combinaison.

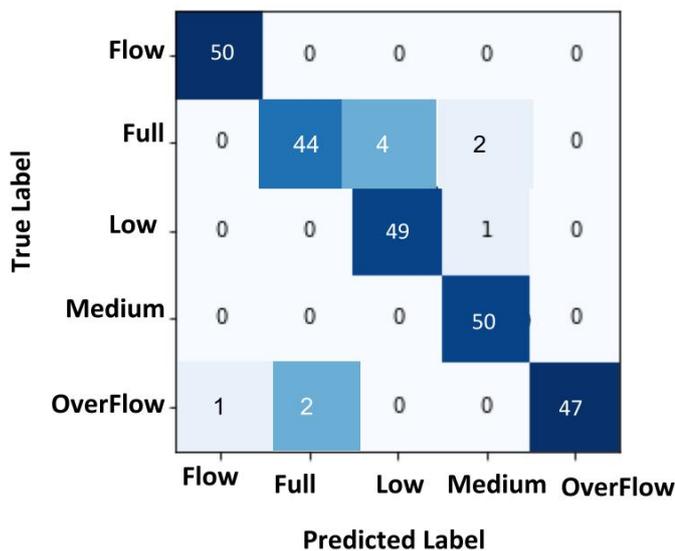


FIGURE 4.13 – Matrice de confusion du 3 plus proches voisins avec la distance Corrélation et la validation croisée (14 partitions)

4.4.2 Résultats du système de recherche adaptative

(a) Résultats de la solution fine-tuning

Dans le tableau 4.19, nous présentons les résultats obtenus pour la technique de fine-tuning avec les 3 modèles VGG-16, MobileNet, ResNet-50. Ce test est effectué en utilisant 25% de la base d'images pour l'évaluation des performances. Nous constatons que les meilleurs résultats sont obtenus par les modèles MobileNet et ResNet-50 avec un F1-score de 98.94% et 98.41%, respectivement. Le modèle VGG-16 donne de très mauvais résultats avec un F1-score de 22%.

Modèle	Précision	Rappel	F1-score
MobileNet	98.94%	98.94%	98.94%
ResNet-50	98.41%	98.41%	98.41%
VGG-16	22%	22%	22%

TABLE 4.19 – Résultats de fine-tuning.

Afin de vérifier la stabilité du classificateur et pour évaluer ses performances, nous allons utilisé la méthode K-Fold pour la validation croisée. Et cela, tout en variant le nombre de partitions de la base.

a) Nombre de partition $k=2$

Nous avons réalisé une validation croisée avec 2 partitions, pour les deux modèles MobileNet et ResNet-50. Le tableau 4.20 présente les résultats des test. Nous constatons essentiellement que la validation croisée a nettement amélioré les résultats. Nous avons eu une précision de 100% pour les deux modèles, nous justifions ceci par la précision qui a déjà été élevée sans validation croisée, de plus la base d'images sur laquelle nous travaillons est considérablement petite.

Modèle	Précision	Rappel	F1-score
MobileNet	100%	100%	100%
ResNet-50	100%	100%	100%

TABLE 4.20 – Résultats de fine-tuning avec la validation croisée.

b) Les Matrices de confusion :

Dans cette partie, nous présentons les différentes matrices de confusion des tests sur le fine-tuning, avec et sans validation croisée.

i. ResNet-50 sans validation croisée

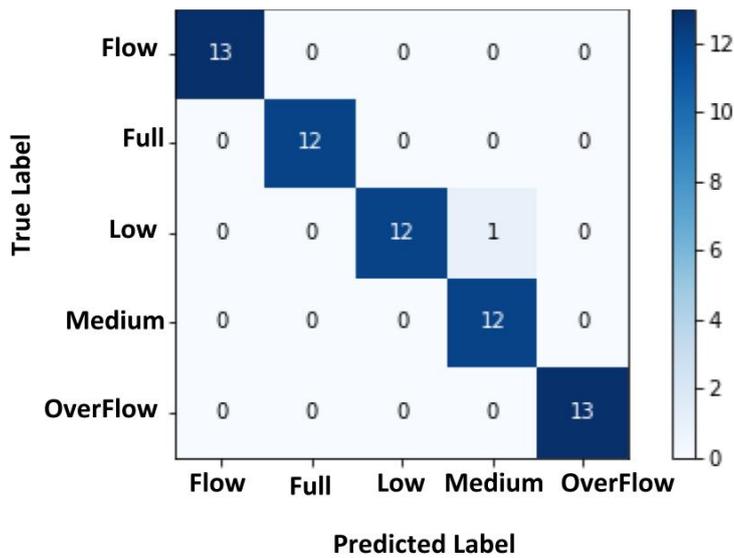


FIGURE 4.14 – Matrice de confusion du modèle ResNet-50 sans validation croisée.

Selon la matrice 4.14, le modèle ResNet-50 avait classé toutes les images de test correctement, mis à part une. Nous observons une confusion entre la classe Low et Medium.

ii. ResNet-50 avec validation croisée

La matrice 4.15 montre l'amélioration de performance avec la validation croisée. Ceci est traduit par la diagonale de la matrice qui prouve que toutes les images ont été correctement classées.

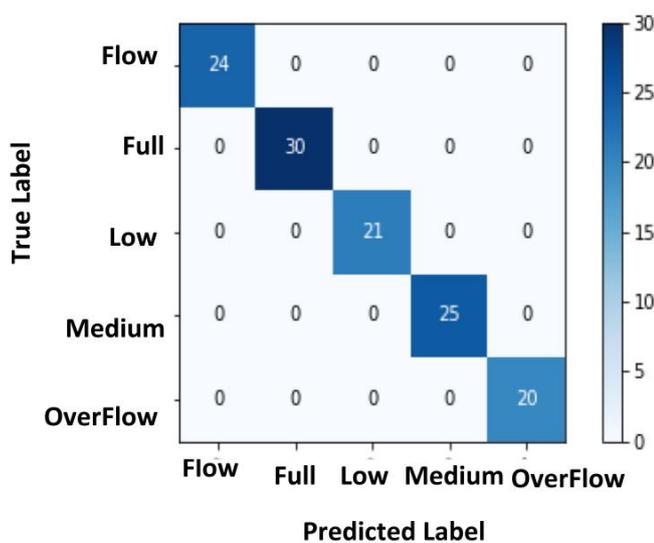


FIGURE 4.15 – Matrice de confusion du modèle ResNet-50 avec validation croisée.

iii. MobileNet sans validation croisée

La matrice de MobileNet 4.16 sans validation croisée, fait référence à deux confusions, la première est entre la classe Flow et Overflow, tandis que la deuxième est entre Low et Medium. Nous constatons que, la matrice 4.17 a été améliorée suite à la validation croisée.

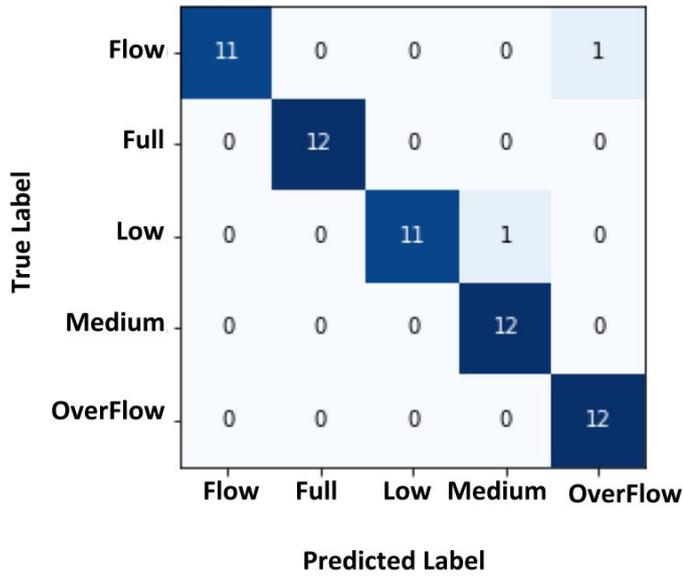


FIGURE 4.16 – Matrice de confusion du modèle MobileNet sans validation croisée.

iv. MobileNet avec validation croisée

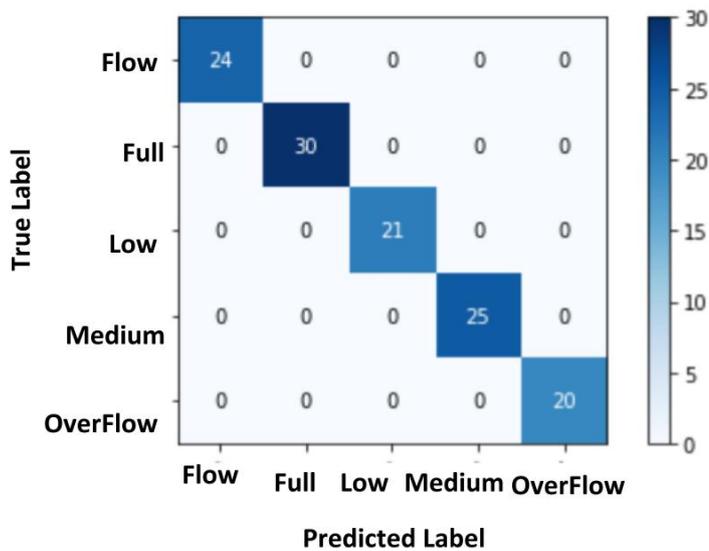


FIGURE 4.17 – Matrice de confusion du modèle MobileNet avec validation croisée.

v. VGG-16 sans validation croisée

Pour le VGG-16, la matrice 4.18, montre plusieurs confusions entre pratiquement toute les classes. Ceci est justifié par la précision très faible.

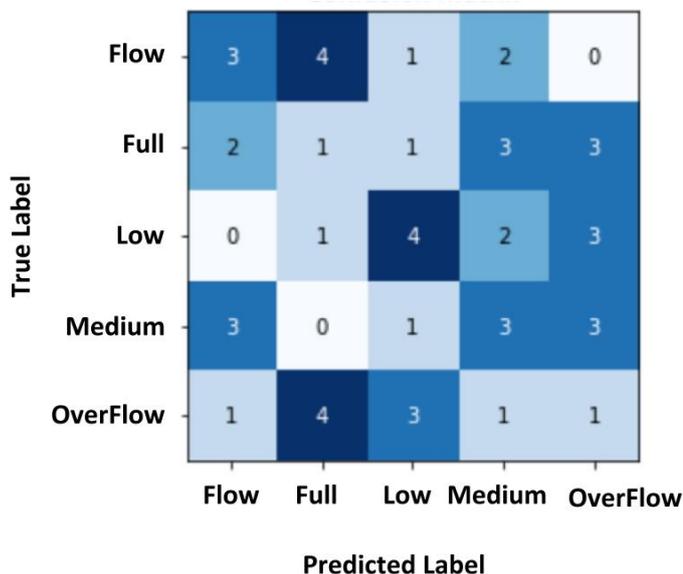


FIGURE 4.18 – Matrice de confusion du modèle VGG-16 sans validation croisée.

(b) **Résultats de la solution basée sur l'apprentissage profond pour l'extraction des caractéristiques**

Dans le Tableau 4.21, nous présentons les résultats obtenus par la technique d'extraction de caractéristiques avec le classificateur SVM, et en se basant sur les trois modèles ResNet-50, MobileNet, et VGG-16. Les tests ont été réalisés sur 25% de la base d'images. Nous remarquons que les résultats sont similaires pour les trois modèles VGG-16-SVM, MobileNet-SVM, ResNet-50-SVM.

Modèle	Précision	Rappel	F1-score
VGG-16 - SVM	98.86%	98.86%	98.86%
MobileNet - SVM	98.86%	98.86%	98.86%
ResNet-50 - SVM	98.86%	98.86%	98.86%

TABLE 4.21 – Résultats du classificateur SVM avec VGG-16, MobileNet, ResNet-50 sans validation croisée.

Dans ce qui suit, nous allons effectuer les tests sur cette technique avec une validation croisée, en variant le nombre de partitions.

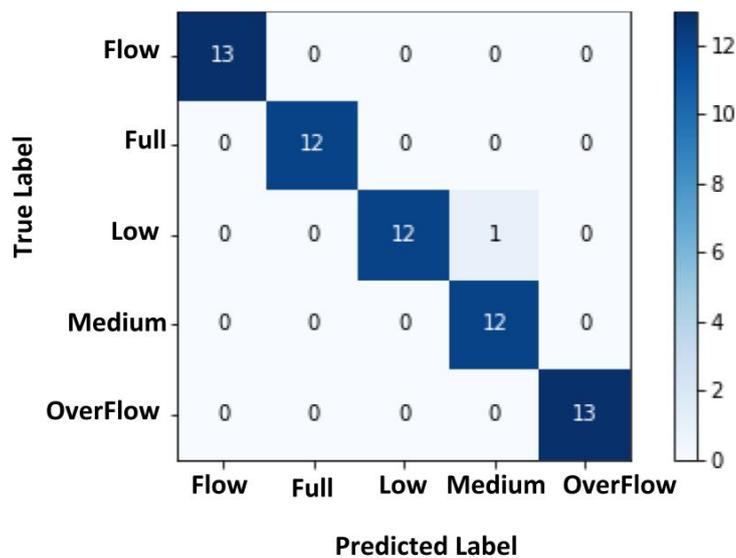


FIGURE 4.19 – Matrice de confusion du modèle SVM sans validation croisée.

a) Nombre de partitions $k=2$

Les résultats suivant, le tableau 4.22, ainsi que la matrice de confusion 4.20 montrent une amélioration considérable du performance des modèles. Nous constatons que la validation croisée augmente le taux de précision et donne un meilleur résultat.

Modèle	Précision	Rappel	F1-score
VGG-16 - SVM	99.61%	99.60%	99.59%
MobileNet - SVM	99.61%	99.60%	99.59%
ResNet-50 - SVM	99.61%	99.60%	99.59%

TABLE 4.22 – Résultats du classificateur SVM avec VGG-16, MobileNet, ResNet-50 avec la validation croisée pour $K=2$.

b) Nombre de partitions $k=3$

Modèle	Précision	Rappel	F1-score
VGG-16 - SVM	99.60%	99.58%	99.58%
MobileNet - SVM	99.60%	99.58%	99.58%
ResNet-50 - SVM	99.60%	99.58%	99.58%

TABLE 4.23 – Résultats du classificateur SVM avec VGG-16, MobileNet, ResNet-50 avec la validation croisée pour $K=3$.

c) Nombre de partitions $k=5$



FIGURE 4.20 – Matrice de confusion du modèle SVM avec validation croisée pour $k=2$.

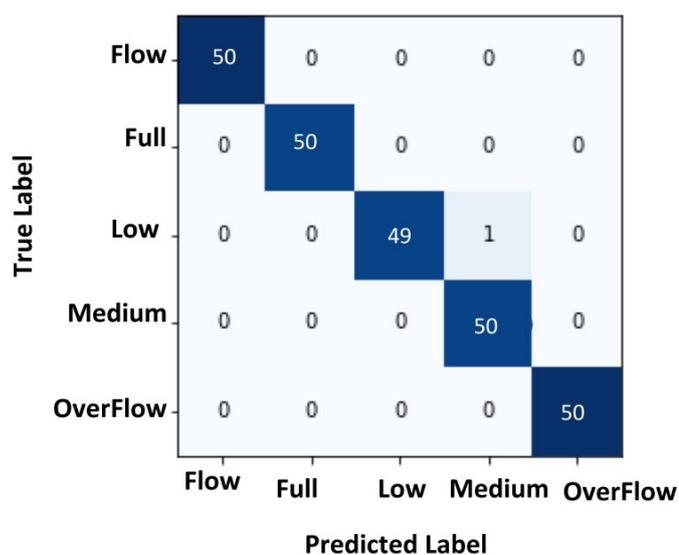


FIGURE 4.21 – Matrice de confusion du modèle SVM avec validation croisée pour $K=3$.

Modèle	Précision	Rappel	F1-score
VGG-16 - SVM	99.58%	99.60%	99.59%
MobileNet - SVM	99.58%	99.60%	99.59%
ResNet-50 - SVM	99.58%	99.60%	99.59%

TABLE 4.24 – Résultats du classificateur SVM avec VGG-16, MobileNet, ResNet-50 avec la validation croisée pour $K=5$.



FIGURE 4.22 – Matrice de confusion du modèle SVM avec validation croisée pour K=5.

d) Nombre de partitions $k=7$

Les résultats du tableau 4.25, donnent les meilleures valeurs de précision, les trois modèles ont aboutit à une performance de 99.64%.

Modèle	Précision	Rappel	F1-score
VGG-16 - SVM	99.64%	99.59%	99.58%
MobileNet - SVM	99.64%	99.59%	99.58%
ResNet-50 - SVM	99.64%	99.59%	99.58%

TABLE 4.25 – Résultats du classificateur SVM avec VGG-16, MobileNet, ResNet-50 avec la validation croisée pour K=7.



FIGURE 4.23 – Matrice de confusion du modèle SVM avec validation croisée pour K=7.

4.4.3 Comparaison avec l'état de l'art

Dans le tableau 4.26 nous comparons les résultats obtenus par rapport à ceux des travaux connexes.

Modèle	Précision	Rappel	F1-score
K-ppv+Correlation	97.10%	95.71	96.40
K-ppv+Cosine	99.07%	98.57%	98.82%
ResNet	100%	100%	100%
VGG-16 + SVM	99.64%	99.59%	99.58%
Gabor- MLP [32]	94%	/	/
MLP [45]	98.5%	/	/
GLCM-MLP [31]	95%	/	/
GLCM-K-ppv [31]	97%	/	/

TABLE 4.26 – Évaluation par rapport aux travaux connexes

Comparés aux résultats rapportés dans [31], [32] et [45], les résultats obtenus par les deux systèmes proposés (l'approche classique et adaptative) sont très intéressants, et surpassent largement les systèmes des travaux connexes. Le système classique a apporté un résultat de 99.07% de précision avec le modèle basé sur le K-ppv en combinaison avec la distance Cosine. Tandis que, le système adaptative a atteint le 100% avec le modèle ResNet. Nous concluons que, la variation des distance de similarité, ainsi que l'architecture profonde des réseaux CNN ont permis aux systèmes d'obtenir d'excellentes performances .

4.5 Conclusion

Dans ce chapitre, nous avons présenté en détail l'évaluation et les tests effectués sur nos solutions. Dans un premier temps, nous avons présenté l'environnement de développement de nos systèmes ainsi que les interfaces pour décrire les différentes fonctionnalités. Dans un second temps, les résultats de la classification ont été bien analysés en utilisant les mesures de performance (Rappel, Précision et F1-score), ces derniers ont trouvé de bons résultats de classification des images.

Conclusion et perspectives

Ce travail de recherche s'inscrit dans le cadre de la gestion des déchets municipaux automatique, ainsi que les techniques de classification d'images. Dans le cadre de ce mémoire, nous avons traité une problématique qui touche le monde entier.

Les déchets constituent un grand défi auquel fait face toutes les sociétés actuelles. Leur évolution et leur diversité menacent l'environnement. Dans les pays en développement, la situation en matière d'élimination des déchets solides est critique, et tend parfois à se dégrader dans un contexte de forte urbanisation. Faute de moyens inadéquats la valorisation des déchets est peu ou pas du tout pratiquée. Le traitement des déchets reste problématique et se traduit par la multiplication des dépôts sauvages qui font l'objet soit d'une mise en décharge non contrôlée soit d'une incinération à l'air libre.

Pour faire face à cette problématique, nous avons adopté une méthodologie de recherche qui consiste à concevoir et implémenter deux grands systèmes de classifications automatiques d'images dont une perspective de gestion des déchets de manière automatique.

En effet, le premier système est basé sur une approche classique qui exploite plusieurs techniques de l'apprentissage automatique, ce système comporte plusieurs modules à savoir :

- Module d'indexation, ce module a pour objectif de pré-traiter et extraire les caractéristiques visuelles de bas niveau de l'image de la corbeille.
- Module de classification, ce module utilise l'algorithme k-ppv pour classer l'image de la corbeille selon les 5 niveaux (Medium, Flow, Full, Low, Overflow).

Quant au deuxième système, il s'appuie sur l'apprentissage profond ainsi que la technologie d'apprentissage par transfert, ce système comporte deux modules à savoir :

- Module d'apprentissage par transfert, ce module a pour but de créer, entraîner et évaluer un modèle classificateur.
- Module de classification, ce module vise à attribuer une classe à l'image de la poubelle à travers du classificateur.

Les systèmes développés dans notre projet ont été évalués par plusieurs critères d'évaluations et plusieurs variantes, les résultats obtenus par les deux systèmes sont encourageants par rapport aux systèmes développés dans la littérature scientifique.

Ainsi, en faisant une analyse sur les résultats obtenus, nous avons constaté que le système basé sur une recherche adaptative a abouti à une performance plus élevée que l'approche classique. En effet, nous avons réussi à concevoir un système basé sur la classification d'images qui permet de reconnaître le taux de remplissage d'une poubelle. Les résultats obtenus après cette classification sont satisfaisants. Nous pouvons donc conclure que la classification d'images avec les méthodes d'apprentissage supervisé profonde est une voie importante de recherche.

Notre travail ouvre des perspectives scientifiques à court et à long terme. Nous soulignons dans ce qui suit les perspectives qui nous semblent pertinentes pour l'évolution des systèmes développés dans ce projet.

- Premièrement, il serait intéressant de travailler avec une base d'images plus large et plus variée.
- Lors de l'aboutissement de ce travail, une deuxième perspective apparaît est de mettre en valeur le temps d'exécution, et de reconnaître le taux de remplissage des poubelles en temps réel.

Bibliographie

- [1] (2018). Convolutional neural networks. <http://cs231n.github.io/convolutional-networks/>.
- [2] (2018). Demystifying optimizations for machine learning. <https://towardsdatascience.com/demystifying-optimizations-for-machine-learning-c6c6405d3eea>.
- [3] (2018). Taking the confusion out of confusion matrices. <https://towardsdatascience.com/taking-the-confusion-out-of-confusion-matrices-clce054b3d3e>.
- [4] (2018). Understanding categorical cross-entropy loss. https://gombru.github.io/2018/05/23/cross_entropy_loss/.
- [5] (2019). About opencv. <https://opencv.org/about/>.
- [6] (2019). Keras : The python deep learning library. <https://keras.io/>.
- [7] (2019). Présentation de matlab. <https://matlabpourtous.com/presentation/presentation-de-matlab/>.
- [8] Akbani, R., Kwek, S., and Japkowicz, N. (2004). Applying Support Vector Machines to Imbalanced Datasets. pages 39–50.
- [9] Alom, M. Z., Taha, T. M., Yakopcic, C., Westberg, S., Sidike, P., Nasrin, M. S., Van Esesn, B. C., Awwal, A. A. S., and Asari, V. K. (2018). The History Began from AlexNet : A Comprehensive Survey on Deep Learning Approaches. *arXiv :1803.01164 [cs]*. arXiv : 1803.01164.
- [10] Arebey, M., Hannan, M. A., Basri, H., Begum, R. A., and Abdullah, H. (2010). Solid waste monitoring system integration based on RFID, GPS and camera. *2010 International Conference on Intelligent and Advanced Systems*, pages 1–5.

- [11] Aude (2007). Mission préparatoire «déchets : découverte du marché algérien».
- [12] Badr Benmammar (2018). *Cours intelligence artificielle et système multi-agents*. Université Abou Bekr Belkaïd.
- [13] Bedouhene, S. (2011). Recherche d'images par le contenu.
- [14] BENSMAIL Salem and AINOUCHE Mohand Cherif (2010). *La problématique de la gestion des déchets solides à travers les modes de traitement des déchets ménagers et hospitaliers : Cas de la commune de Bejaia*. PhD thesis, ABDERRAHMANE MIRA DE BEJAIA.
- [15] Candillier, L. (2011). Contextualisation, Visualisation et Evaluation en Apprentissage Non Supervisé.
- [16] Caragliu, A., Del Bo, C., and Nijkamp, P. (2009). Smart Cities in Europe. *VU University Amsterdam, Faculty of Economics, Business Administration and Econometrics, Serie Research Memoranda*, 18.
- [17] Celisse, A. and Mary-huard, T. (2011). Exact Cross-Validation for k NN : application to passive and active learning in classification. 152(3) :83–97.
- [18] cognitive and data science (2017). Multilayer perceptron. <http://srutisj.in/2017-06-15-MultiLayerPerceptron/>.
- [19] Dertat, A. (2017). Applied Deep Learning - Part 4 : Convolutional Neural Networks.
- [20] Dipanjan Sarkar, R. B. and Ghosh, T. (2018). *Hands-On Transfer Learning with Python : Implement advanced deep learning and neural network models using TensorFlow and Keras*. Packt Publishing, 1 edition edition.
- [21] Dr. B. L. Malleswari and Mrs. P. Nalini (2016). Review on Content Based Image Retrieval : From Its Origin to the New Age. 3.
- [22] Dr.BENNAMA Tahar (2016). *Les bases de traitement des déchets solides*. Université des Sciences et de la Technologie d'Oran « Mohamed Boudiaf » Faculté de Chimie – Département de Génie Chimique.
- [23] Echopédia (2014). Collecte sélective des déchets. https://www.ekopedia.fr/wiki/Tri_sélectif.
- [24] F. Russo, C. Rindone, and P. Panuccio (2014). The process of smart city definition at an EU level. *The Sustainable City*, 2.
- [25] focus, A. (2016). Recyclage des déchets ménagers et assimilés/ une pratique qui peine à s'installer en algérie. <https://www.algerie-focus.com/2016/06/recyclage-dechets-menagers-assimiles-pratique-peine-a-sinstaller-algerie>

- [26] Fukushima, K. (1988). Neocognitron : A hierarchical neural network capable of visual pattern recognition. *Neural Networks*, 1 :119–130.
- [27] Georgios Drakos (2018). Cross Validation.
- [28] Gueguen, L. (2013). Thèse e : Signal et Images. (January 2007).
- [29] Hackeling, G. (2014). *Mastering Machine Learning With scikit-learn*. Packt Publishing - ebooks Account, Birmingham.
- [30] Hamza CHENITI (2014). *LA GESTION DES DECHETS URBAINS SOLIDES : CAS DE LA VILLE D'ANNABA*. PhD thesis, Université Badji Mokhtar-Annaba-.
- [31] Hannan, M. A., Arebey, M., Begum, R. A., Basri, H., and Al Mamun, M. A. (2016). Content-based image retrieval system for solid waste bin level detection and performance evaluation. *Waste Management*, 50 :10–19.
- [32] Hannan, M. A., Arebey, M., Begum, R. A., Mustafa, A., and Basri, H. (2013). An automated solid waste bin level detection system using Gabor wavelet filters and multi-layer perception. *Resources, Conservation and Recycling*, 72 :33–42.
- [33] Hassan CHOUAIB (2011). *Sélection de caractéristiques : méthodes et applications*. PhD thesis, l'université Paris Descartes.
- [34] He, K., Zhang, X., Ren, S., and Sun, J. (2015). Deep Residual Learning for Image Recognition. *arXiv :1512.03385 [cs]*. arXiv : 1512.03385.
- [35] Hiba, B. (2016). Système d'analyse des sentiments.
- [36] Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M., and Adam, H. (2017). MobileNets : Efficient Convolutional Neural Networks for Mobile Vision Applications. *arXiv :1704.04861 [cs]*. arXiv : 1704.04861.
- [37] Jadon, S. (2018). Introduction to different activation functions for deep learning. <https://medium.com/@shrutijadon10104776/survey-on-activation-functions-for-deep-learning-9689331ba092>.
- [38] Jason Brownlee (2019). How to Configure the Learning Rate Hyperparameter When Training Deep Learning Neural Networks.
- [39] Lecun, Y., Bottou, L., Bengio, Y., and Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11) :2278–2324.
- [40] Lee, C. K. M., Yeung, C. L., Xiong, Z. R., and Chung, S. H. (2016). A mathematical model for municipal solid waste management – A case study in Hong Kong. *Waste Management*, 58 :430–441.

- [41] Manaf, L., Abu Samah, M. A., and Ilyana Mohd Zukki, N. (2009). Municipal solid waste management in Malaysia : Practices and challenges. *Waste management (New York, N.Y.)*, 29 :2902–6.
- [42] MathWorks (2019a). Classement à l'aide de plus proches voisins. <https://fr.mathworks.com/help/stats/classification-using-nearest-neighbors.html>.
- [43] MathWorks (2019b). Entropy. <https://fr.mathworks.com/help/images/ref/entropy.html>.
- [44] Maïtine Bergounioux (2011). *Quelques méthodes de filtrage en Traitement d'Image*.
- [45] Md. Shafiqul Islam, M.A. Hannan, Hassan Basri, Aini Hussain, and Maher Arebey (2014). Solid waste bin detection and classification using Dynamic Time Warping and MLP classifier. *Waste Management*, 34(2) :281–290.
- [46] Mezouri Sandjakdine (2011). *Conception et exploitation des centres de stockage des déchets en Algérie et limitation des impacts environnementaux*. PhD thesis, école polytechnique d'architecture et d'urbanisme.
- [47] Mitran, M. (2014). *Annotation d'images via leur contexte spatio-temporel et les métadonnées du Web*. thesis, Toulouse 3.
- [48] Olivas, E. S., Guerrero, J. D. M., Sober, M. M., Benedito, J. R. M., and Lopez, A. J. S. (2009). *Handbook Of Research On Machine Learning Applications and Trends : Algorithms, Methods and Techniques*. Information Science Reference, Hershey, PA, 1 edition edition.
- [49] Quellec, G. (2008). Indexation et fusion multimodale pour la recherche d'information par le contenu. application aux bases de données d'images médicales.
- [50] RAPHAËL, M. J. B. (2017). Machine learning : Introduction à l'apprentissage automatique. <https://www.supinfo.com/articles/single/6041-machine-learning-introduction-apprentissage-automatique>.
- [51] Rohith Gandhi (2018). A Look at Gradient Descent and RMSprop Optimizers.
- [52] ROSENBLATT, F. (1957). *The Perceptron : A Perceiving and Recognizing Automaton (Project PARA)*. Report No. 85-460-1.
- [53] Sarkar, D. (2018). A comprehensive hands-on guide to transfer learning with real-world applications in deep learning.
- [54] Simonyan, K. and Zisserman, A. (2014). Very Deep Convolutional Networks for Large-Scale Image Recognition. *arXiv :1409.1556 [cs]*.

- [55] soir d'algerie, L. (2018). Le volume des déchets va doubler d'ici 2035. <https://www.lesoirdalgerie.com/actualites/le-volume-des-dechets-va-doubler-dici-2035-15407>.
- [56] Tani, S. M. (2016). Indexation et Recherche d' image médicale à partir de la transformée en ondelette : Applications aux images rétinienne.
- [57] The world bank and METAP (2004). projet régional de gestion des déchets solides dans les pays du Mashreq et Maghreb. rapport du pays-Algérie.
- [58] theffork (2019). activation functions. <https://theffork.com/activation-functions-in-neural-networks/>.
- [59] Torrey, L. and Shavlik, J. (2010). Transfer Learning. *Handbook of Research on Machine Learning Applications and Trends : Algorithms, Methods, and Techniques*, pages 242–264.
- [60] Walia, A. S. (2017). Types of optimization algorithms used in neural networks and ways to optimize gradient descent.
- [61] world Bank, T. (2019). Solid waste management. <https://www.worldbank.org/en/topic/urbandevelopment/brief/solid-waste-management>.