



جامعة العربي التبسي - تبسة
Université Larbi Tébessi - Tébessa

République Algérienne Démocratique et Populaire
Ministère de l'enseignement supérieur et de la
recherche scientifique

Université Larbi Tébessi - Tébessa

Faculté des Sciences Exactes et des Sciences de la Nature et de la Vie
Département : Mathématiques et Informatique



كلية العلوم الدقيقة وعلوم الطبيعة والحياة
FACULTÉ DES SCIENCES EXACTES
ET DES SCIENCES DE LA NATURE ET DE LA VIE

Mémoire de fin d'études
En vue de l'obtention du diplôme de **MASTER**
Domaine : Mathématiques et Informatique
Filière : Informatique
Option : Systèmes et Multimédia

Thème

Repérage des mots dans les documents manuscrits off-ligne

Présenté Par :
Nebba Abdelghani

Devant le jury :

Dr. Abdeljalil Gattal	MCB	Université Larbi Tébessi - Tébessa	Président
Dr. Bouakkez Fatima	MAA	Université Larbi Tébessi - Tébessa	Examineur
Dr. Faycel Abbas	MAA	Université Larbi Tébessi - Tébessa	Encadreur
Dr. Yousfi Douaâ		Université Larbi Tébessi - Tébessa	Co-Encadreur

Date de soutenance : **Juin 2022**



Université Larbi Tébessi- Tébessa

Faculté des sciences exactes et des sciences de la nature et de la vie

Département de *Maths et informatique*

Filière : *Informatique*

Année universitaire 2021/2022



Déclaration sur l'honneur de non-plagiat (A joindre obligatoirement avec le mémoire)

Je, soussigné(e)

Nom et prénom :

Nebba Abdelghani

Régulièrement inscrit (e) :

N de carte d'étudiant :

98603102

Année universitaire :

2021/2022

Domaine :

Mathématiques et informatique

Filière :

Informatique

Spécialité :

Systèmes et multimedia

Intitulé :

Reperage des mots dans les documents off-line.

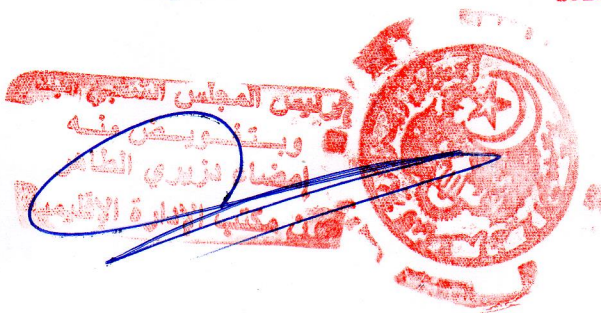
Atteste que mon mémoire est un travail original et que toutes les sources utilisées ont été indiquées dans leur totalité, je certifie également que je n'ai ni copié ni utilisé des idées ou des formulations tirées d'un ouvrage, article ou mémoire, en version imprimée ou électronique, sans mentionner précisément leur origine et que les citations intégrales sont signalées entre guillemets.

Sanctions en cas de plagiat prouvé :

L'étudiant sera convoqué devant le conseil de discipline, les sanctions prévues selon la gravité de plagiat sont :

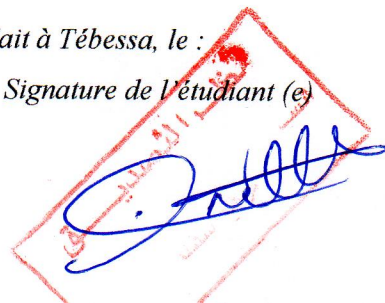
- L'annulation du mémoire avec possibilité de refaire sur un sujet différent.
- L'exclusion d'une année de Master.
- L'exclusion définitive.

2022 جوان 25



Fait à Tébessa, le :

Signature de l'étudiant (e)



Remerciement

Nous remercions **ALLAH** Tout-Puissant, d'abord et avant tout, de nous avoir donné la force et patience et surmonter les difficultés. pour terminer ce que nous avons commencé dans le voyage d'étude . Nous demandons à Dieu de nous aider dans ce qui reste de cette vie

Je tiens à exprimer mes profonds remerciements à mon encadreur **Dr. Faycel Abbaset** Co-Encadreur **Dr. Yousfi Douaâ** , pour les encadrement, la confiance et pour le soutien qu'ils m'ont apportés durant tout ce travail.

Leurs compétences scientifiques ont toujours été sources d'enrichissement qui me permettant de faire à bien ce travail. Merci pour tous vos conseils et votre patience.

Un grand merci à tous les membres du Jury qui me font l'honneur d'accepter de juger mon travail.

Je remercie tous les enseignants et les étudiants du département d'informatique de l'Université de Tébessa,.

Je voudrais aussi présenter mes remerciements à ma mère et mes frères pour tous leurs sacrifices en faveur de nos éducations, mes amis pour leur support qui j'ai été autorisé à de travailler dans les meilleures conditions.

Résumé :

Le repérage des mots dans les documents manuscrits modernes est un problème intrigant, car il permet la recherche, l'indexation et la classification d'images de documents. En conséquence, cela a été un problème de domaine de recherche actif, avec des progrès significatifs réalisés ces dernières années. Malgré des progrès significatifs, le repérage des mots manuscrits peut encore à peine fonctionner de manière acceptable sur des images de documents manuscrits dont le style et la qualité d'écriture varient considérablement.

Cette mémoire présente une nouvelle méthode de repérage des mots dans des documents manuscrits dans un cadre basé sur la segmentation où des caractéristiques extraites d'images de mots sont utilisées pour faire matching un mot-clé de requête avec ceux d'une base de référence. Plus précisément, nous utilisons des caractéristiques d'image de base orientées (**oBIFs**) pour caractériser les images de mots, et la matching est effectuée en calculant la distance entre les images de mots à l'aide de la distance **City-block**.

Une étude expérimentale du système est réalisée à l'aide de la base de données de la compétition de repérage des mots **ICFHR 2014** et des résultats prometteurs sont rapportés en termes de précision **P@K** et de **meanaverageprecision(mAP)**.

Mots-clés: Le repérage des mots, Documents manuscrits modernes, Caractéristiques d'image de base orientées (**oBIFs**), Distance **City-block**.

Abstract:

Word spotting in modern handwritten documents is an intriguing problem because it enables document image search, indexing, and classification. As a result, this has been an active research area problem, with significant progress made in recent years. Despite significant progress, handwritten word spotting can still barely perform acceptably on images of handwritten documents that vary greatly in style and quality of writing.

This thesis presents a new method for word spotting in handwritten documents in a segmentation-based framework where features extracted from word images are employed to match a query keyword with those in a reference base. More specifically, we use oriented basic image features (**oBIFs**) to characterize the word images, and matching is performed by computing the distance between word images using **City-block** distance.

An experimental study of the system is carried out using the dataset of the **ICFHR 2014** word spotting competition and promising results are reported in terms of **P@K** precision and mean average precision (**mAP**).

Keywords: Word spotting, Modern Handwritten documents, Oriented Basic Image Features (**oBIFs**), **City-block** distance.

خلاصة :

يعد اكتشاف الكلمات في المستندات المكتوبة بخط اليد الحديثة مشكلة مثيرة للاهتمام لأنه يتيح البحث في صورة المستند وفهرستها وتصنيفها. ونتيجة لذلك، كانت هذه مشكلة نشطة في مجال البحث، مع إحراز تقدم كبير في السنوات الأخيرة. على الرغم من التقدم الكبير، فإن اكتشاف الكلمات المكتوبة بخط اليد لا يزال بالكاد يؤدي بشكل مقبول على صور المستندات المكتوبة بخط اليد والتي تختلف اختلافاً كبيراً في أسلوب وجودة الكتابة. تقدم هذه الأطروحة طريقة جديدة لاكتشاف الكلمات في المستندات المكتوبة بخط اليد في إطار عمل قائم على التجزئة حيث يتم استخدام الميزات المستخرجة من صور الكلمات لمطابقة الكلمات الرئيسية للاستعلام مع تلك الموجودة في قاعدة مرجعية. وبشكل أكثر تحديداً، نستخدم ميزات الصورة الأساسية الموجهة (oBIFs) لتوصيف صور الكلمات، ويتم إجراء المطابقة عن طريق حساب المسافة بين صور الكلمات باستخدام مسافة كتلة المدينة. تم إجراء دراسة تجريبية للنظام باستخدام مجموعة البيانات الخاصة بمنافسة اكتشاف الكلمات ICFHR 2014 وتم الإبلاغ عن نتائج واعدة من حيث دقة $P @ K$ ومتوسط الدقة . (mAP)

الكلمات الرئيسية: اكتشاف الكلمات، الحديثة المكتوبة بخط اليد، وميزات الصورة الأساسية الموجهة (oBIFs)، ومسافة كتلة المدينة.

TABLE DES MATIÈRES

	Page
Remerciements	
Resumés	
Introduction générale	12
Chapitre I : Introduction a repérage des mots	
Introduction	15
1.1. Définition de repérage des mots (word sptting)	15
1.2. Différents aspects de repérage des mots	16
1.2.2. Les méthodes de repérage des mots	16
1.2.2.1. La méthode globale (sans ségmentation)	16
1.2.2.2. La méthode analytique (basé sur ségmentation)	16
1.2.3. Les catégories de repérage des mots	16
1.2.3.1. Le repérage des mots basé sur une requête image (Query-by-Exemple)	16
1.2.3.2. Le repérage des mots basé sur une requête textuelle (Quert-by-String)	16
1.3. Procédure generale de repérage des mots	16
1.3.1. Prétraitement	17
1.3.1.1. La binarisation	18
1.3.1.2. Le lissage	18
1.3.1.3. La squelettisation	19
1.3.1.4. La suppression de bruit	19
1.3.1.5. La normalisation	20
1.3.2. Segmentation	20
1.3.2.1. La segmentation en lignes	21
1.3.2.2. La segmentation en mot (pseudo-mot)	21
1.3.2.3. La segmentation en caractères	22
1.3.3. Extraction de caractéristique	22
1.3.3.1. Modèle Binarie Local (LBP)	22
1.3.3.2. Histogramme des Gradients Orientés (HOG)	23
1.3.4. Matching	23
1.3.4.1. Déformation Temporelle Dynamique (DTW)	23
1.3.4.2. Les classificateurs de distance	24
1.3.4.2.1. Distance Corrélacion	24
1.3.4.2.2. Distance Cosinus	24
1.3.4.2.3. Distance Euclidienne	24
Conclusion	25
Chapitre II : Un état de l'art	
Introduction	27
2.1. Les travaux ulteriers	27

2.1.1. Rothfeder et al et al. 2003	27
2.1.2. Lladós et al. 2012	27
2.1.3. Aldavert et al. 2013	28
2.1.4. Almazán et al. 2014	28
2.1.5. Kovalchuk et al. 2014	28
2.1.6. Yao et al. 2015	29
2.1.7. Retsinas et al et al. 2018	29
2.1.8. Le compétition de repérage de mots-clés manuscrits 2014 (H-KWS 2014)	30
2.2. Comparaison des travaux	30
Conclusion	31
Chapitre III : Méthode proposes et résultats expérimentaux	
Introduction	33
3.1. Méthode proposes pour le repérage des mots	33
3.1.1. Le système proposée	33
3.1.2. Présentation de la basse de donnees	34
3.1.3. Extraction de caracteristique	34
3.1.3.1. Caractéristiques d'image de base orientées (oBIFs)	35
3.1.4. Matching	36
3.2. Présentation des outils de développement	36
3.2.1. Matériel	36
3.2.2. Matlab	37
3.3. Résultats et comparaison	37
3.3.1. Les métriques utilisées	37
3.3.2. Les résultats expérimentaux	37
3.4. Comparaison entre les travaux de compétitions et la méthode proposée	40
Conclusion	41
Conclusion générale	43
Référence bibliographie	48

LISTE DES FIGURES

N°	Titre	Page
Figure 1.1	Principe d'un système de repérage des mots	17
Figure 1.2	Exemple de binarisation	18
Figure 1.3	Résultat de lissage	19
Figure 1.4	La squelettisation de l'image	19
Figure 1.5	La suppression de bruit	20
Figure 1.6	La segmentation en lignes	21
Figure 1.7	La segmentation en mot	21
Figure 1.8	La segmentation en caractères	22
Figure 3.1	Le système proposée	33
Figure 3.2	Échantillons d'images de documents manuscrits de la base de données	34
Figure 3.3	Exemple de calcul d'oBIFs pour une image de mot manuscrit pour $\sigma=2$ et $\sigma=4$ Alors que $\varepsilon=0.001$	36
Figure 3.4	Échantillons des mots de requête sélectionnés à partir de la base de données Moderne	38
Figure 3.5	Performance de la méthode proposée à l'aide des oBIFs	40

LISTE DES TABLEAUX

N°	Titre	Page
Table 2.1	Comparaison des travaux	30
Table 3.1	Comparaison entre que le paramètre $\varepsilon= 0.1$, $\varepsilon= 0.01$ et $\varepsilon= 0.001$	38
Table 3.2	Performances avec différents paramètres dans le calcul des oBIFs	39
Table 3.3	Performance de la méthode proposée en fonction de différentes métriques de distance	40
Table 3.4	Performance de comparaison entre les travaux de compétitions et la méthode proposée	41

LISTE DES ABRÉVIATIONS

OCR	Optical Character Recognition
WS	Word Spotting
QbE	Query-by-Example
QbS	Query-by-String
DTW	Dynamic Time Warping
LBP	Local Binary Patterns
HOG	Histogram of Oriented Gradients
BIF	Basic Image Feature
oBIFs	OrientedBasic Image Feature
SVM	Support Vector Machin
POG	Projection of Oriented Gradients
SIFT	Scale-Invariant Feature Transform
SURF	Speeded Up Robust Features
SSD	Sum of Squared Differences
BoVW	Bag of Visual Words
GW	George Washington
LB	Lord Byron
mAP	mean Average Precision

INTRODUCTION GÉNÉRALE

Introduction générale

Grâce aux progrès des technologies de l'information et de la communication, ces dernières années ont connu une croissance spectaculaire de la quantité de documents manuscrits conservés, traités et consultés sous forme numérique. Les documents historiques en tant que sous-ensemble de documents manuscrits sont les précieuses ressources pour les chercheurs afin que leur contenu puisse être mis à disposition via Internet ou d'autres médias électroniques. Le problème principal est que ces contenus ne sont disponibles qu'en format d'image, ce qui les rend difficiles à rechercher. Le repérage des mots est devenu une technique très efficace pour la recherche d'informations et l'indexation à grande échelle d'ensembles de données de documents manuscrits.

Le repérage des mots (Word spotting : WS) a été proposé comme alternative à la transcription complète pour récupérer des mots-clés à partir d'images de documents. Un système de repérage des motstypique comprend un ensemble de documents (base de référence) et, lorsqu'il est fourni avec un élément de requête, le système produit une collection de documents qui sont pertinents pour la requête. Dans cette perspective, les systèmes de repérage des mots sont similaires aux systèmes traditionnels de recherche d'informations.

Le repérage des mots consiste à récupérer toutes les instances d'un mot de requête donné à partir de collections d'images de documents (généralement manuscrites) à l'aide de techniques de matching d'images sans l'intervention d'un OCR.

En fonction du type de requête d'entrée, les méthodes de détection de mots divisées en deux catégories, Query-by-Example (QbE) et Query-by-String (QbS). Dans l'approche QbE, l'utilisateur fournit une image du mot sous forme de requête, tandis que l'approche QbS permet à l'utilisateur de fournir la requête sous forme de texte. Dans de nombreux cas, la chaîne de requête fournie est convertie en une image utilisée dans le cadre de QbE.

D'autre part, d'un point de vue méthodologique, les techniques de repérage des mots se distinguent en approches basées sur la segmentation et sans segmentation. Dans les approches basées sur la segmentation, les mots interrogés peuvent être directement comparés aux images de mots segmentés, tandis que les techniques sans segmentation reposent sur la comparaison de parties de documents pour correspondre à la requête sans aucune segmentation en entités sémantiquement significatives (mots par exemple).

Un système de repérage des mots typique comprend les composants clés du prétraitement, de la segmentation, de l'extraction des caractéristiques et de le matching; l'extraction de caractéristiques et les étapes de matching étant au centre de l'étude actuelle. L'extraction de caractéristiques est la partie la plus critique de tout système de matching de modèles. Des représentations d'entités efficaces peuvent conduire à des performances élevées, même avec des techniques de matching simples. Cette étude examine l'efficacité des caractéristiques d'image de base orientées (oBIFs) pour le repérage des mots manuscrits. Les caractéristiques extraites des mots manuscrits segmentés sont comparées à celles des images de requête à l'aide d'un certain nombre de mesures de distance: la distance de DTW, la corrélation, Spearman, Cosinus, Euclidean, Chebychev, Cityblock. En fonction du degré de similarité, les documents contenant le mot-clé requêtes sont retournés à l'utilisateur.

Objectifs de mon travail :

- ❖ Etat de l'art sur différentes méthodes de repérage des mots.
- ❖ L'objectif du travail est l'évaluation des performances du système de repérage des mots dans les documents manuscrits moderne en utilisant les caractéristiques d'image de base orientées (oBIFs). L'évaluation est réalisée sur la base de données ICFHR 2014 moderne.

Organisation du mémoire:

Pour faire, ce travail se doit sous deux aspects, l'une théorique, l'autre applicatif. Il sera présenté relativement selon les chapitres suivants :

Le premier chapitre contient une généralité sur le repérage des mots, puis définit le processus d'un système de repérage des mots et présentation de certaines techniques d'extraction des caractéristiques (**LBP, oBIFs, HOG...**) et le matching (**Correlation, Euclidean, City block, Chebychev, Cosinus, Spearman, ...**).

Le deuxième chapitre est présenté les différents travaux pour le repérage des mots.

Dans le dernier chapitre, un système basé sur les caractéristiques d'image de base orientées (oBIFs) et City-block est présenté, où un ensemble d'expérimentation sont faites sur la base de données **ICFHR 2014 moderne**.

Enfin, nous terminerons ce mémoire par une conclusion générale et perspective pour les futurs travaux.

CHAPITRE I:

INTRODUCTION A REPÉRAGE DES MOTS

Introduction

Ces dernières années, le repérage des mots est devenu une technique très efficace pour la recherche d'informations et l'indexation de bases de données de documents manuscrits à grande échelle. Le repérage des mots a été proposé comme alternative à la reconnaissance optique de caractère (OCR)[1] pour récupérer les mots-clés des images de documents.

Dans ce chapitre, nous passerons en revue certains concepts et définitions fondamentaux liés au repérage des mots.

1.1. Définition de repérage des mots (Word spotting)

Le repérage des mots est une technique qui permet de rechercher plusieurs occurrences d'une requête d'entrée en utilisant des méthodes de matching dans toutes les images de document. Il permet également la création d'index partiels pour les collections de documents traités similaires aux index de fin de livre. Ainsi, le repérage des mots facilite l'indexation et la recherche d'informations données sous forme de requête dans des documents historiques et modernes numérisés relativement complexes et dégradés.

Les étapes de recherche de la requête sont :

Premièrement, elles construisent l'index basés sur des caractéristiques de bas niveau extraites d'images de mots. Les mêmes fonctionnalités sont également extraites de la requête images. Deuxièmement, ils consultent l'ensemble de la base de données et mesurent combien l'image de requête est similaire à chaque image de mot dans la base de données. Enfin, les résultats sont triés en fonction du niveau de matching de similarité.

Le repérage des mots a été initialement proposé par Manmatha, dans [Manmatha, et al., 1996] et [Manmatha, et al., 1996b], puis il a incité un certain nombre de publications qui proposent des algorithmes et des caractéristiques comme l'approche [Rath, et al., 2007], [Adamek, et al., 2007], [Rothfeder, et al., 2003], [Rath, et al., 2003].

1.2. Diffèrent aspects de repérage des mots

1.2.1. Les méthodes de repérage des mots

1.2.1.1. Laméthode globale (sans segmentation)

Cette méthode prend un aperçu du mot et s'appuie sur une description unique de l'image du mot en tant qu'entité indivisible. Cette méthode a l'avantage de préserver le caractère dans son contexte environnant, permettant une modélisation plus efficace des changements et dégradations de l'écriture qu'elle peut subir[2].

1.2.1.2. Laméthodeanalytique (basé sur segmentation)

Basées sur la segmentation d'un mot, une image de document ou une image de mot est segmentée en unités plus petites qui peuvent être reconnues indépendamment ou après regroupements partiels[3].

1.2.2. Les catégorie du repérage des mots

1.2.2.1. Le repérage des mots base sur une requête image (Query-by-Exmple)

Le repérage des mots basé sur une requête image(QbE) ces approches consistent à retrouver dans les documents[4], es images de mots similaires à l'image de requête donnée par l'utilisateur.

1.2.2.2. Le repérage des mots base sur une requête textuelle (Query-by-String)

Le repérage des mots basé sur une requête textuelle (QbS) ces approches permettent à l'utilisateur de fournir sa requête sous la forme de texte pour rechercher dans les documents, les images de mots correspondantes[5].

1.3. Procédure générale de repérage des mots

Le principe de repérage des mots permet de rechercher d'un mot dans une base de données, en trouvant les meilleures matching entre les caractéristiques de la requête et les caractéristiques des images de la base de données.

Un système de repérage des mots fait appel généralement aux étapes suivantes : prétraitement, segmentation, extraction des caractéristiques et leur matching. Le schéma suivant explique ce système (**Figure 1.1**):

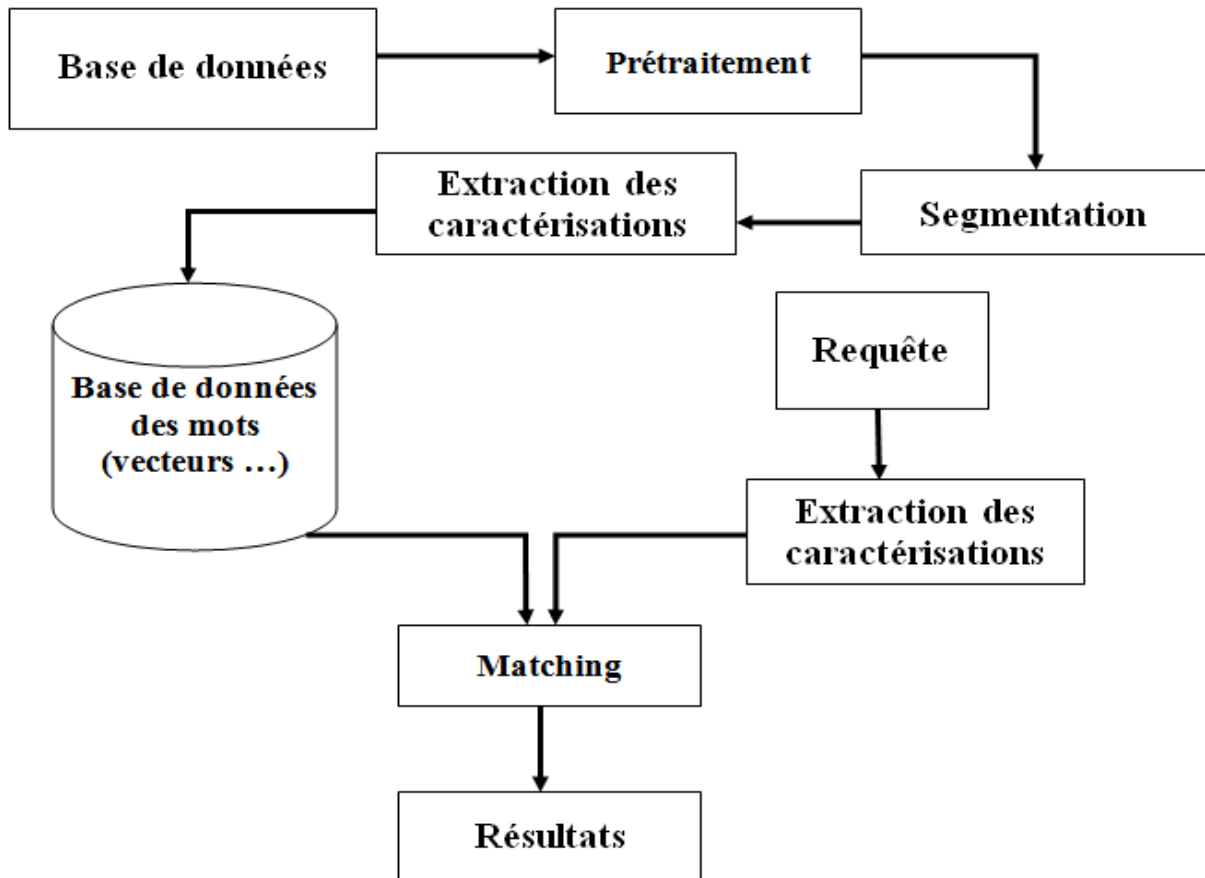


Figure 1.1 | Principe d'un système de repérage des mots

1.3.1. Prétraitement

Avant de commencer la reconnaissance de mots, un certain nombre de pré-traitements doivent être effectués sur l'image afin de supprimer (réduire)[6] le bruit qui l'engorge et de simplifier les étapes d'extraction de caractéristiques et segmentation. Les techniques de pré-traitement comprennent [7] :

1.3.1.1. La binarisation

La binarisation permet de passer à une image binaire composée de deux valeurs 0 et 1. Pour cerner et traduire les contrastes dans l'image. Un seuil de binarisation est défini. Cependant, Si dans l'image la distribution de niveaux de gris n'est pas clairement bimodale, il est difficile de fixer un seuil. La plus simple des façons de le mesurer est de calculer l'histogramme des niveaux de gris de l'image [8]. Dans ce cas, la valeur du seuil est égale à la valeur du niveau de gris entre les deux pics de l'histogramme. La détermination du seuil peut être locale ou globale[9], les méthodes locales sont déterministes, mais elles sont plus coûteuses en termes de temps de calcul contrairement aux méthodes globales qui ne donnent pas d'exactes résultats, mais en temps plus réduit [10].(Figure 1.2)

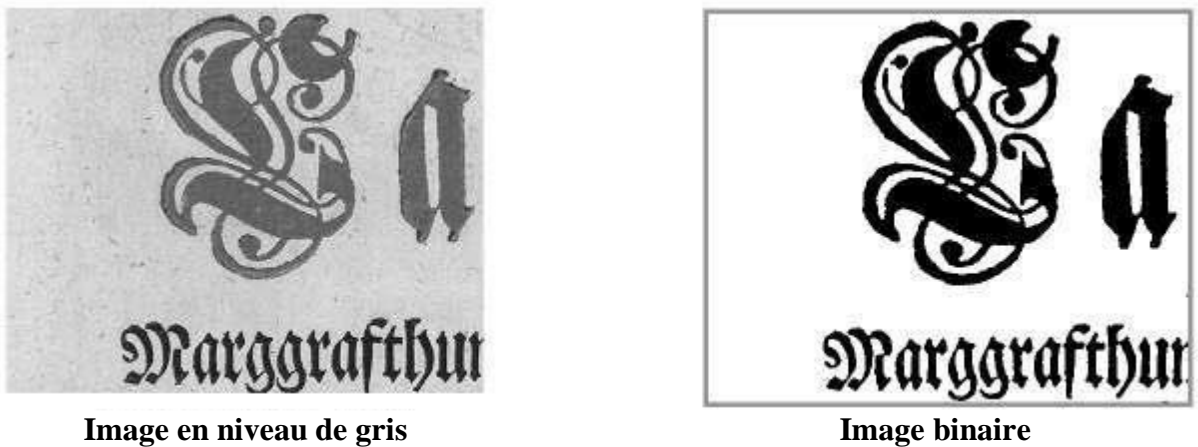


Figure 1.2| Exemple de binarisation[11]

1.3.1.2. Le lissage

Durant le processus de l'acquisition, l'image est sujette à différents types de bruit. Elle peut avoir des absences de pixels ou une surcharge. Pour corriger cela, nous appliquons le lissage sur les images bruitées. C'est l'analyse du voisinage des pixels pour éliminer les pixels isolés pour appliquer le lissage consiste à parcourir l'image pixel par pixel. L'analyse de chaque pixel se fait en se basant sur ses 8 voisins.

Les pixels de valeur 1 sont mis à 0 s'il n'y a pas assez de pixels noirs autour d'eux et les pixels de valeur 0 sont mis à 1 s'il n'y a pas assez de pixels blancs autour d'eux. Une autre approche consiste à inverser la valeur du pixel (entre 0 et 1) si la somme des pixels voisins est inférieure à un seuil déterminé[12].(Figure 1.3)



Image avec résultat de binarisation



Image traitée

Figure 1.3| Résultat de lissage[11]

1.3.1.3. La squelettisation

Pour applique la squelettisation sur l'image il faut doit image binaire, l'image binaire est une succession de traits d'épaisseurs variables, généralement avec bruit Par conséquent, la squelettisation s'applique sur les images après la binarisation, le but de la squelettisation est de réduire l'épaisseur du tracé des caractères à un pixel, en concevant sa continuité. La procédure consiste à faire des opérations d'érosion conditionnelle successive jusqu'à ce que le but soit atteint[13]. Le squelette obtenu doit préserver la forme du caractère, et ne doit pas introduire du bruit[14].(Figure 1.4)



Figure 1.4| La squelettisation de l'image [15]

1.3.1.4. La suppression de bruit

Lors de la phase d'acquisition de l'image, de nouveaux composants peuvent s'introduire à l'image : c'est ce qu'on appelle bruit, le prétraitement consiste à détecter les pixels qui n'appartiennent pas aux composants significatifs de l'image et à les supprimer. Cela permet d'augmenter la discrimination de nouveau l'image. [16].(Figure 1.5)



Image en niveau de gris



Image filtré

Figure 1.5| La suppression de bruit [17]

1.3.1.5. La normalisation

La normalisation est appliquée sur les images pour réduire tous les types de variations, et pour obtenir des données normalisées. Cependant, il peut engendrer une distorsion et élimine quelques informations utiles. Usuellement, la normalisation porte sur la taille des caractères, l'inclinaison des lignes et l'inclinaison des caractères.

Les méthodes existantes pour normalisation sont :

- Normalisation de la taille.
- Correction de l'inclinaison des lignes (Skew correction).
- Correction de l'inclinaison des caractères (Slant correction).
- Estimation de la ligne de base.

1.3.2. Segmentation

La segmentation est une étape cruciale car elle extrait les régions significatives pour l'analyse, qui permet de découper l'image du texte en entité (phrase, mot, caractère), tout dépend de l'objectif du système. L'entité de sortie est toujours une matrice de pixels.

L'étape de segmentation peut être divisée en trois niveaux : segmentation en lignes, segmentation en mots (pseudo-mots) et segmentation en caractères :

1.3.2.1. La segmentation en ligne

Dans la segmentation du texte en lignes, les lignes sont extraites à l'aide la méthode de la projection horizontale, où la méthode calcule l'histogramme horizontal de l'image[6]. Si le nombre de lignes blanches successives rencontrées (du haut vers le bas) est supérieur à un seuil fixé. Le nombre de ligne est augmenté, puis le système extrait le début et la fin de chaque ligne. (Figure1.6)

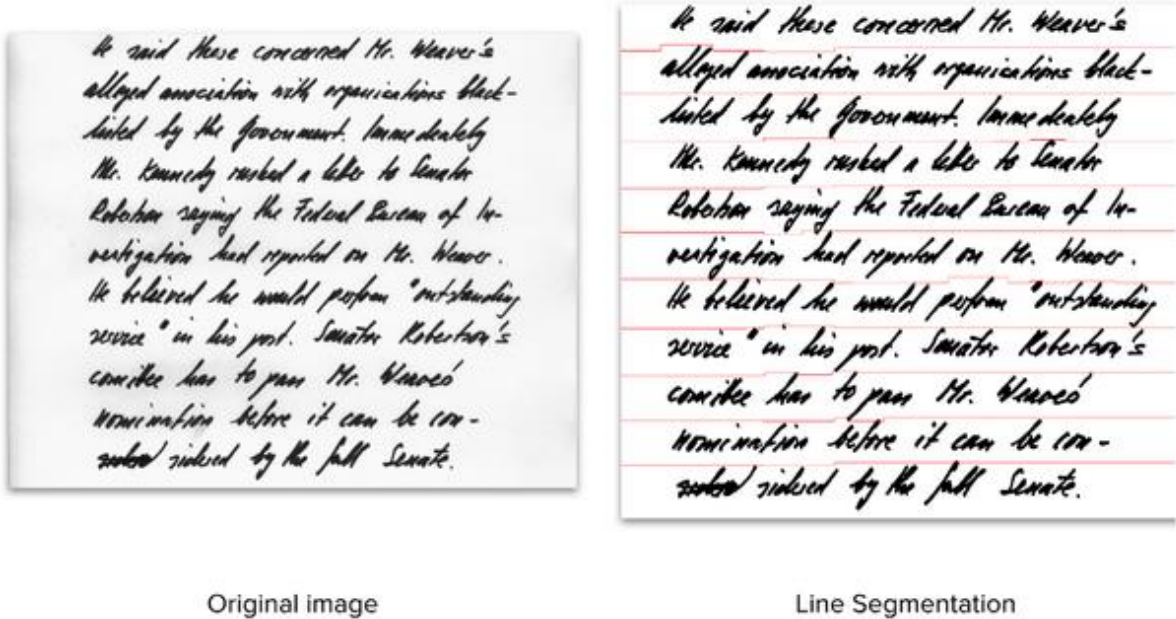
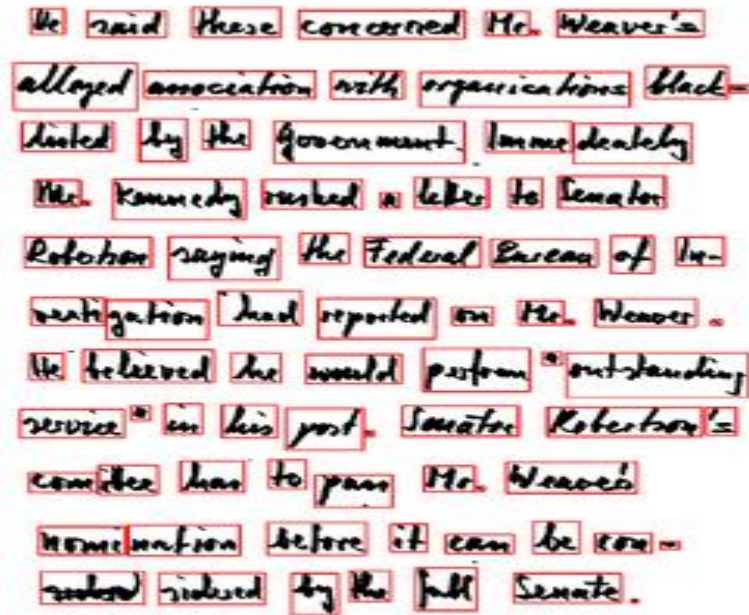


Figure 1.6| La segmentation en lignes[18]

1.3.2.2. La segmentation en mot(pseudo-mot)

La segmentation en mot (**pseudo-mot**)cette étape suit la segmentation en lignes de texte. Il est pour l'analyse de la ligne pour la découper en mots (pseudo-mots). Cette étape pour réaliser en déterminant l'histogramme des projections verticales des lignes pour séparer entre les mots avec détecter les espaces entre eux. Pour chaque colonne de l'image texte, il faut compter le nombre de pixels noirs. Si le nombre de pixels noirs n'est pas égal à zéro, il indique une partie connectée et par conséquent le texte n'est pas segmenté. D'autre part, si le nombre de colonnes blanches successives rencontré est supérieur à un seuil, la partie courante est segmentée. (Figure 1.7)



Word Segmentation

Figure 1.7| La segmentation en mot[18]

1.3.2.3. La segmentation en caractères

La segmentation des caractères[6] être après L'étape du processus de prétraitements qui permet de séparer les lignes dans une structure de données définie puis séparer les caractères de chaque ligne. L'histogramme est la technique choisie pour cette segmentation. L'histogramme de projection horizontale permet de segmenter les lignes et l'histogramme de projection vertical sur l'image d'une ligne permet de segmenter les caractères de cette ligne.(Figure 1.8)

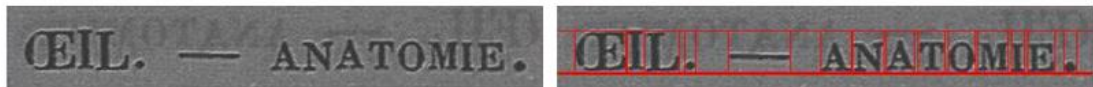


Figure 1.8| La segmentation en caractère [19]

1.3.3. Extraction de caractéristique

Le choix d'une représentation des caractéristique robuste est l'étape la plus importante de tout problème de matching de modèles, car l'étape d'extraction des caractéristiques réduit la dimension des images originales tout en préservant et en extrayant les informations importantes encodées dans l'image. L'objectif principal de l'extraction de caractéristiques est de représenter

les images (de mots) de manière que les objets (mots) ayant des propriétés morphologiques similaires aient tendance à se regrouper.

Il existe plusieurs méthodes d'extraction des caractéristiques comme : Modèle Binaire Local (**LBP**), Histogramme des Gradients Orientés (**HOG**)...

1.3.3.1. Modèle Binaire Local(LBP)

LBP a été proposé initialement par **Ojala et Al(2007)** pour décrire les textures (caractériser la texture) présentes dans les images en niveaux de gris. Est un opérateur, très simple et très efficace. Le calcul de la valeur **LBP** consiste pour chaque pixel à seuiliser ses huit voisins directs avec un seuil dont la valeur est le niveau de gris du pixel courant. Sa taille est de **3 x 3** pixels. Tous les voisins prendront alors une valeur **1** si leur valeur est supérieure ou égale au pixel courant et **0** si leur valeur est inférieure. Le masque local pour l'extraction des codes LBP est déplacé dans les directions horizontale et verticale. [20].

1.3.3.2. Histogramme des Gradients Orientés (HOG)

Les descripteurs **HOG (Histogram of Oriented Gradients)** ont été introduits par **Dalal et Triggs dans [21] et [22]**. L'idée principale derrière l'histogramme de gradient est que l'apparence locale et la forme de l'objet dans un l'image peut être décrite par la distribution d'intensité des gradients ou la direction des contours. La mise en œuvre de ces les descripteurs peut être obtenus en divisant l'image en petits régions connectées, appelées cellules. Ensuite, pour chaque cellule, nous calculons un histogramme des directions de gradient ou des orientations de bord pour tous pixels de la cellule. La combinaison de ces histogrammes est le descripteur.

Le descripteur **HOG** présente certains avantages clés. Puisqu'il opère sur des cellules localisées, la méthode maintient l'invariance aux transformations géométriques et photométriques.

1.3.4. Matching

L'étape de matching est au cœur du système de repérage des mots. Il effectue la recherche et en comparant le vecteur de caractéristiques d'images de requête au vecteur de caractéristiques d'images de tous les mots de référence. Le but de ce processus est de déterminer les images de mots des documents qui sont similaires à une image de mots de requête donnée en calculant la similarité des vecteurs caractéristiques extraits [23].

La similarité (distance) entre deux vecteurs de caractéristiques est calculée à l'aide de certain nombre des mesures : la distance de DTW, les classificateurs de distance (la corrélation ,Cosinus, Euclidean, City-block,...).

1.3.4.1. Déformation Temporelle Dynamique (DTW)

La déformation temporelle dynamique (algorithme DTW) est un algorithme permettant de mesurer la similarité entre deux suites qui peuvent varier au cours du temps. Il a été exploité en vidéo, audio, graphique, Bio-informatique et dans autre domaine. Et il peut être appliqué dans toute situation où les données peuvent être transformées en une représentation linéaire. Ce qu'il le met attractive dans notre situation. D'une façon générale, la DTW est une méthode qui recherche un matching optimal entre deux séries temporelles, sous certaines restrictions. Les séries temporelles sont déformées par transformation non-linéaire de la variable temporelle, pour déterminer une mesure de leur dis similarité, indépendamment de certaines transformations non-linéaires du temps[24].

1.3.4.2. Les différents types de distance

1.3.4.2.1. Distance Correlation :

La corrélation de distance est une mesure de la force d'association entre des variables aléatoires non-linéaires. Il va au-delà de la corrélation de Pearson, car il peut détecter plus que des associations linéaires et il peut fonctionner de manière multidimensionnelle. La corrélation de distance varie de 0 à 1, où 0 implique l'indépendance entre X et Y et 1 implique que les sous-espaces linéaires de X et Y sont égaux[25].

La formule de corrélation de distance est la suivante :

$$dCor(X, Y) = \frac{dCor^2(X, Y)}{\sqrt{dVar(X)dVar(Y)}} \dots (1)$$

1.3.4.2.2. Distance Cosinus :

La similarité cosinus est une métrique qui peut être utilisée pour déterminer à quel point les objets de données sont similaires, quelle que soit leur taille. Les objets de données dans une base de données sont traités comme des vecteurs en similarité cosinus[26].

La formule pour calculer la similarité cosinus de deux vecteurs est :

$$\mathit{Cos}(x, y) = \frac{x \cdot y}{||x|| * ||y||} \quad \dots (2)$$

3.4.3.2.3. Distance Euclidienne :

La métrique de distance euclidienne a été initialement utilisée par CANDID. Il l'a utilisé pour la comparaison des signatures globales[27]. La distance euclidienne peut être calculée comme suit : Si A(x,y) et B(p,q) sont deux pixels, alors la distance euclidienne peut être donnée par :

$$\mathit{distance} = \sqrt{(|x - p|^2 + |y - q|^2)} \quad \dots (3)$$

3.4.3.2.4. Distance City-block :

La valeur de la distance du City-block est toujours supérieure ou égale à 0. Nous obtiendrons 0 s'il y a des points identiques sous considérations. La valeur sera élevée s'il y a une faible similarité. La distance du City-block peut être calculée comme suit : si A (x, y) et B (p, q) sont deux pixels, alors la distance du City-block peut être donnée comme suit :

$$\mathit{distance} = |x - p| + |y - q| \quad \dots (4)$$

Conclusion

Dans ce chapitre, nous avons présenté les différentes phases d'un système de repérage des mots, Ainsi, nous avons démontré les problèmes liés à la complexité de la reconnaissance. Nous présentons dans le chapitre suivant, une description d'un état de l'art et les travaux ultérieurs et la comparaison des travaux.

CHAPITRE II:
UN ÉTAT DE L'ART

Introduction

Ce chapitre est dédié à la description de l'état de l'art dans le domaine de repérage des mots. Nous nous concentrons dans ses premières sections sur la présentation des principaux travaux de recherche dans le domaine.

Ensuite, nous terminons le chapitre par une comparaison des différents travaux du domaine en utilisant, comme critères de comparaison, la base de données utilisée, la méthode d'extractions des caractéristiques choisies, la méthode de matching et le mAP.

2.1.Les travaux ultérieurs

Le repérage des mots est resté un domaine de recherche attractif pour la communauté de la classification des modèles, à la fois pour les documents manuscrits et imprimés. Parmi l'un des travaux notables sur ce sujet :

2.1.1.Rothfeder et al et al. 2003

Parmi d'autres méthodes connues qui ont rapporté des performances de récupération prometteuses. Une méthode statistique basée sur le détecteur de coins de Harris(Harris Corner Detector) est présentée par **Rothfeder et al [28]**. Pour comparer des images de mots segmentés. Le matching entre les points angulaires est utilisée pour classer les images de mots par similarité. Le matching entre ces points sont établies en comparant les fenêtres locales et en utilisant la somme des carrés des différences (SSD). Des expériences ont été menées sur deux bases de données de documents historiques avec 2372 (bonne qualité) et 3262 (mauvaise qualité) images et ont rapporté une meanaverageprecision(mAP) de 62,5% et 15,5% respectivement.

2.1.2.Llados et al. 2012

Dans **Llados et al[29]**, ont réalisé une étude comparative pour évaluer la performance de différents descripteurs pour repérer des mots-clés dans des documents historiques en utilisant un cadre de requête par exemple. Le matching basée sur le déformation temporelle dynamique (DTW) est utilisée comme ligne de base tandis que d'autres fonctionnalités incluent un sac de mots visuels (BoVW), un modèle pseudo-structurel et une approche structurelle qui représente les mots sous forme de graphiques. Des expériences ont été menées sur deux ensembles de données,

George Washington (GW) et Cathédrale de Barcelone et ont rapporté une meanaverageprecision(mAP) de 42.20 % et 30.00% respectivement, une étude expérimentale a révélé que les caractéristiques du sac de mots (statistiques) surpassent les autres représentations.

2.1.3.Aldavert et al. 2013

Une autre étude de **Aldavert et al[30]**. Proposent un cadre de requête par chaîne où les images de mots dans l'ensemble d'apprentissage sont représentées à la fois par des représentations textuelles et visuelles. La représentation textuelle est basée sur des n-grammes de caractères tandis que la représentation visuelle repose sur le schéma du sac de mots visuels. Les deux représentations sont fusionnées et projetées dans un sous-espace vectoriel en utilisant la méthode d'analyse sémantique latente. Cette transformation a permis, à partir d'une requête textuelle, de récupérer des instances de mots qui n'étaient représentés que par la modalité visuelle. Des expériences ont été menées sur deux ensembles de données, tous les mots en tant que requêtes de George Washington(GW) et les requêtes de vocabulaire de George Washington(GW) et ont rapporté une meanaverageprecision(mAP) de 56,54 % et 76,20 % respectivement.

2.1.4.Almazán et al.2014

La méthode proposée par **Almazán et al[31]**, évite la segmentation en utilisant une approche de fenêtre glissante pour localiser les régions de document les plus similaires à la requête. Les descripteurs HOG et le frameworkExemplar (SVM) sont utilisés pour représenter les documents et la requête de manière supervisée. La méthode proposée est évaluée à l'aide de deux bases de données, de George Washington (GW) et Lord Byron (LB). Les résultats rapportés étaient de 38,28 % et 78,01 % en termes de mAP pour les deux ensembles de données.

2.1.5. Kovalchuk et al. 2014

Kovalchuk et al[32]. Présente une méthode est basée sur l'extraction des caractéristiques HOG et LBP, après que chaque image de mot est redimensionnée dans un rectangle fixe, résultant en un très grand descripteur. Un opérateur de similarité cosinus et une mise en commun maximale sont utilisés pour réduire la dimensionnait des caractéristiques à un descripteur plutôt

petit (250 D). La récupération est effectuée en classant les mots cible par rapport à leur distance euclidienne.

Cette méthode est étudiée sur la base de données de George Washington (GW) et Lord Byron (LB), le taux de meanaverageprecision(mAP) obtenu respectivement est 50.10% et 90.70%.

2.1.6. Yao et al. 2015

Yao et al[33] ont conçu une méthode de repérage des mots dans laquelle une déformation temporelle dynamique bidirectionnelle est appliquée aux descripteurs HOG. Ici, les descripteurs HOG détectés le long de la même ligne sont normalisés en tant que vecteur, et les descripteurs détectés le long de la même colonne sont normalisés en tant qu'un autre vecteur. En outre, un procédé DTW bidirectionnel est appliqué aux vecteurs normalisés pour déterminer la distance entre l'image utilisée pour l'interrogation et les vecteurs normalisés trouvés à partir des descripteurs HOG. Les résultats sont affichés sur la base de données de George Washington (GW) dans un contexte sans segmentation. Et le résultat a été $mAP = 57.2 \%$.

2.1.7. Retsinas et al. 2018

Dans cet article **Retsinas et al [34]**, une méthode de repérage des mots basée sur la segmentation sur des documents manuscrits est proposée. La méthode consiste en trois étapes, à savoir le prétraitement, l'extraction de caractéristiques et le matching. Au cours de l'étape d'extraction de caractéristiques, une séquence de descripteurs est générée à l'aide d'une combinaison d'un schéma de zonage et d'un nouveau descripteur d'apparence, appelés projections modifiées de gradients orientés (POG). De plus, un schéma de zonage inégal est introduit en appliquant un zonage plus dense uniquement sur interrogez les images pour une représentation plus détaillée. Cela conduit à une réduction significative des besoins de stockage d'une collection de documents. La distance entre la requête et les séquences de mots est efficacement calculée par l'algorithme de matching sélectif proposé. L'efficacité de la méthode proposée est démontrée par une expérimentation menée sur la base de données de Bentham et Modern et le taux de meanaverageprecision(mAP) obtenu respectivement est 70.2 % et 46.8 %.

2.1.8. Le compétition de repérage de mots-clés manuscrits 2014(H-KWS 2014)

Le plus notable d'entre eux est le compétition de repérage de mots-clés manuscrits 2014[35] en collaboration avec la Conférence internationale sur les frontières de la reconnaissance de l'écriture manuscrite (ICFHR). Le concours comprenait deux tâches, le repérage basé sur la segmentation et le repérage sans segmentation. Parmi les systèmes participants, les précisions moyennes les plus élevées pour les deux tâches étaient respectivement de 52.3 % et 74.4 % sur les ensembles de données Modern et Bentham.

2.2. Comparaison des travaux

Un résumé des techniques notables basées sur le repérage des mots rapportées dans la littérature est présenté dans le **Table 2.1**. On peut voir que bien que la performance des différentes études varie en fonction de la technique ainsi que de l'ensemble de données utilisé, en général, la (mAP) sont faibles et offrent un potentiel substantiel à la communauté de recherche pour trouver des solutions robustes à ce problème difficile.

Méthodes	Extraction des caractéristique	Maching	Basede données	mAP (%)
Rothfeder et al. 2003 [28]	Harris Corner Detector	Sum of Squared Differences	Documents istorique 2372(bonne qualité)	62.5
			3262 (mauvaise qualité)	15.5
Llados et al. 2012 [29]	Bag of Visual Words	Distance Cosine	GW	42.2
			Cathédrale de Barcelone	30
Aldavert et al. 2013 [30]	Textualdescriptoret Bag of Visual Words	Distance Cosine	GW	56.54
			Vocabulaire GW	76.2
Almazán et al. 2014[31]	Histogram of Oriented Gradients	Support Vector Machin	GW	33.28
			LB	78.01
Kovalchuk et al. 2014 [32]	Local Binary PatternsetHistogram of Oriented Gradients	Distance Euclidien	GW	50.1
			LB	90.7
Yao et al. 2015	Histogram of Oriented	Dynamic Time	GW	57.2

[33]	Gradients	Warping		
Retsinas et al. 2018 [34]	Projection of Oriented Gradients	Algorithme de matching Sélectif	Bentham	70.2
			Modren	46.8
H-KWS 2014 [35]	Local Binary Patternset Histogram of Oriented Gradients	Distance Euclidien	Bentham	52.4
			Modren	33.8
	Scale-Invariant Feature Transform	Support Vector Machin	Bentham	74.4
			Modren	52.3
	Scale-Invariant FeatureTransformet Speeded Up Robust Features	Distance Euclidien	Bentham	46.2
			Modren	27.8

Table 2.1 | Comparaison des travaux

Conclusion

Dans ce chapitre, nous avons présenté un aperçu sur l'état de l'art dans le domaine de repérage des mots. Dans la première partie, nous nous concentrons sur la présentation des principaux travaux de recherche dans ce domaine et discutons les diverses compétitions. Ensuite, nous terminons par une comparaison des différents travaux du domaine en nous basant, comme critères de comparaison, les bases de données utilisées, le chapitre suivant se consacre à la description l'architecture du **oBIFs** pour caractériser les images de mots choisis et le résultat expérimental.

CHAPITRE III

MÉTHODE PROPOSÉ ET RÉSULTATS EXPÉRIMENTAUX

Introduction

Ce chapitre a été alloué pour présenter les résultats obtenus pour la validation de la méthode proposes de repérage des mots. Nous faire la conception d'application en utilisant le langage MATLAB.

Dans ce chapitre, nous utilisons les caractéristiques d'image de base orientées (**oBIFs**) pour l'Extraction de caractéristique et **City-block** pour faire le Matching. Cette étude expérimentale du système est réalisée à l'aide de base de données **ICFHR 2014**.

3.1. Méthode proposes pour le repérage des mots

3.1.1. Le système proposée

La méthode proposée dans la présente étude repose sur une approche basée sur la segmentation où l'emplacement des mots dans les images du document est fourni dans la vérité terrain. La technique repose sur deux étapes : l'extraction des caractéristiques des images de mots et le Matching. Un aperçu graphique de la méthode est présenté à **la Figure 3.1**. Tandis que les détails de l'extraction et le matching sont présentés dans les sections suivantes.

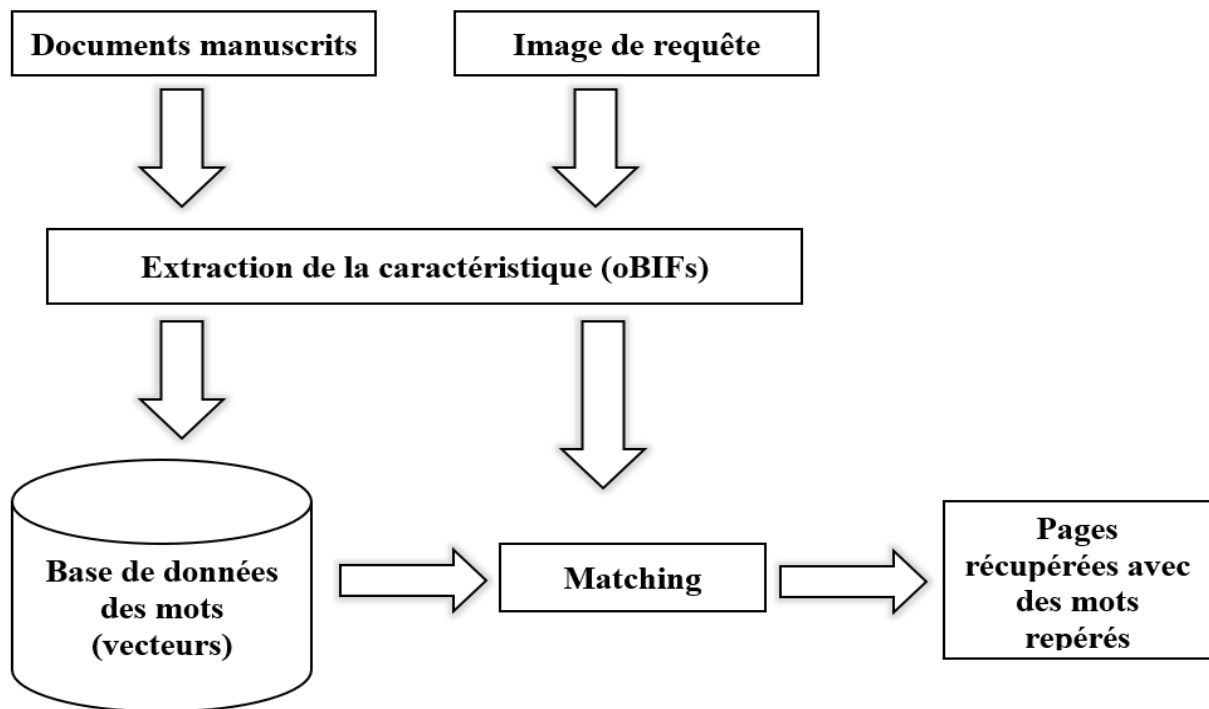


Figure 3.1| Le système proposée[44]

3.1.2. Présentation de la basse de données

L'étude expérimentale de la méthode proposée est réalisée sur la base de données **ICFHR 2014** de compétition de repérage de mots-clés manuscrits (**H-KWS 2014**) [35] pour la piste basée sur la segmentation où les emplacements des images de mots dans le document sont fournis. La base de données se compose de documents manuscrits du concours de segmentation manuscrite **ICDAR 2009** [36], rédigés en quatre langues (anglais, français, allemand et grec). La base de référence compte au total **100** images de documents (**25** pour chacune des quatre langues), et **300** images de mots de requête. Des exemples d'images de documents de la base de données sont illustrés à la **Figure 3.2**.

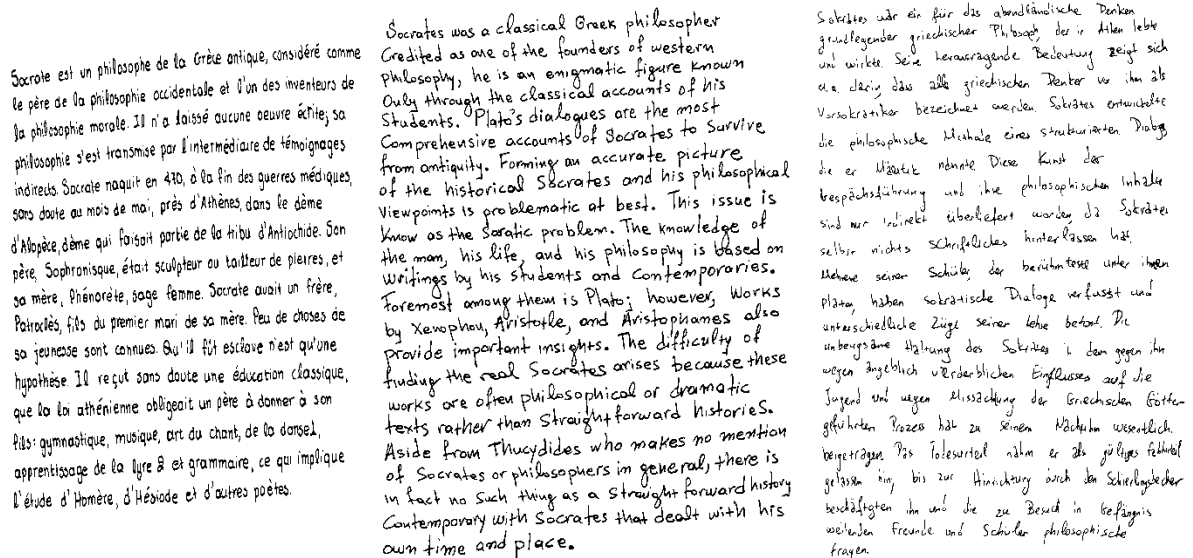


Figure3.2| Échantillons d'images de documents manuscrits de la basede données.

3.1.3. Extraction de caractéristique

Le choix d'une extraction de caractéristiques est l'étape la plus importante dans tout problème de matching. Dans la première série d'expériences, nous avons choisi la méthode caractéristique d'image de base orientées (**oBIFs**) pour d'extraction de caractéristique.

Cette méthode a été soumise par Andrew Newell et Lewis Griffin du Département d'informatique, Collège universitaire de Londres, Royaume-Uni.

3.1.3.1. Caractéristiques d'image de base orientées(oBIFs)

Les **oBIFs**[37,38] ont été appliqués avec succès à un certain nombre de problèmes de classification de modèles tels que la classification de texture[37], la reconnaissance de chiffres [39], l'identification de l'écrivain [37], [40],[41], la classification par sexe [42] et identification de la personnalité basée sur l'écriture manuscrite [43].

Le descripteur **oBIFs** est une extension des caractéristiques d'image de base (BIF) [37,38].Où le calcul consiste à étiqueter chaque emplacement de l'image dans l'une des sept classes de symétrie locales. Les types de symétrie incluent plat, pente, rotation sombre, rotation claire, ligne sombre sur la lumière, ligne claire sur le sombre ou en forme de selle. La réponse d'un banc de six filtres dérivés de Gauss (jusqu'au second ordre) de taille est déterminée par le paramètre d'échelle σ . Le paramètre ε détermine si un emplacement doit être classé comme plat. L'orientation locale pouvant être attribuée à chaque emplacement de l'image dépend du type de symétrie locale. Si l'emplacement est attribué à la classe de rotation claire, de rotation sombre ou à plat, aucune orientation n'est attribuée. Un total de n orientations possibles peut être attribué à la ligne sombre sur les classes claire, la ligne claire sur les classes sombres et en forme de selle, tandis que la classe de pente comprend un total de $2n$ orientations possibles.Il en résulte un vecteur caractéristique de dimension $5n+3$.

Dans notre implémentation des **oBIFs**, nous fixons le paramètre de quantification d'orientation à $n=4$. Cela donne un total de **23 entrées**, l'impact du paramètre d'échelle σ ($\sigma \in \{1,2,4,8,16\}$), et le paramètre ε est fixé à une petite valeur de **0,001**presque, il a donné un bon résultat meilleur que $\varepsilon = 0,1$ et $\varepsilon = 0,01$.

Du point de vue du problème de repérage des mots, nous encodons tous les mots en utilisant la représentation **oBIFs** qui est ensuite utilisée pour le matching.



Figure3.3| Exemple de calcul d'oBIFs pour une image de mot manuscrit pour $\sigma=2$ et $\sigma=4$ Alorque $\epsilon=0.001$. [44]

3.1.4. Matching

Le Matching consiste à comparer une image de mot de requête avec celles de la base de référence à des fins de récupération. Étant donné un mot de requête, nous calculons sa représentation **oBIFs** et le vecteur de caractéristiques résultant est comparé à ceux de tous les mots de référence. La distance (similarité) entre deux vecteurs de caractéristiques est calculée à l'aide de la distance d'une **City-block**[45], tandis qu'un certain nombre d'autres mesures sont également étudiées. Celles-ci incluent la distance **Euclidienne**, la distance **Cosinus**, la distance **Corrélation**, distance **Spearman** et la distance **Chebychev**.

3.2. Présentation des outils de développement

3.2.1 Matériel

Le matériel réalisé est PC personnels :

- Processeur : Intel® Core™ i3-2120U CPU @ 3.30GHz 3.30 GHz.
- Mémoire (RAM) : 4.00 GB.
- Type de système : Système d'exploitation Windows 7 édition intégral, service pack 1, 32 bits.

3.2.2. Matlab

LeMATLAB est la contraction matrix laboratory développé par la société **The MathWorks** [46]. Nous avons utilise dans ma projet **MATLAB 2013a** .

Le MATLAB est devenu un langage de référence pour l'analyse et la résolution de problème scientifique [47]. Il intégré à la fois des solutions de calcul, de visualisation et un environnement de développement [48]. Ce logiciel est spécialement conçu pour le calcul scientifique et la manipulation de vecteurs, et l'objet le plus commun dans Matlab est la matrice.

3.3. Résultats et comparaison

3.3.1. Les métriques utilisées

Pour quantifier les performances du système dans différentes expériences, nous utilisons les métriques standard, notamment precision à k ($P@k$) et mean Average Precision (**mAP**).

Ces métriques sont élaborées dans ce qui suit :

➤ $P@k$ la précision à k est obtenue en calculant la précision à un rang de coupures donné, en ne considérant que les k premiers résultats renvoyés par le système [49]. Dans nos expériences, nous rapportons les résultats pour trois valeurs de k (**1, 5 et 10**).

➤ Formellement, la mesure $P@k$ est définie comme suit :

$$P@k = \frac{|\{\text{mots pertinents}\} \cap \{k \text{ mots récupérés}\}|}{|\{k \text{ mots récupérés}\}|} \dots (5)$$

➤ Le meanAveragePrecision(**mAP**) pour un ensemble donné d'images de requête est la moyenne des scores de précision moyens pour chaque requête.

La précision moyenne d'une requête est définie comme suit :

$$AP = \frac{\sum(p(k)*rel(k)nk=1}{|\{\text{mots récupérés}\}|} \dots (6)$$

où $rel(k)$ prend la valeur **1** si un mot de rang k est pertinent et la valeur **0** sinon.

3.3.2. Les résultats expérimentaux :

Nous avons mené une série d'expériences pour évaluer l'efficacité des fonctionnalités oBIF pour repérer les mots dans les images de documents la **Figure. 3.3.**

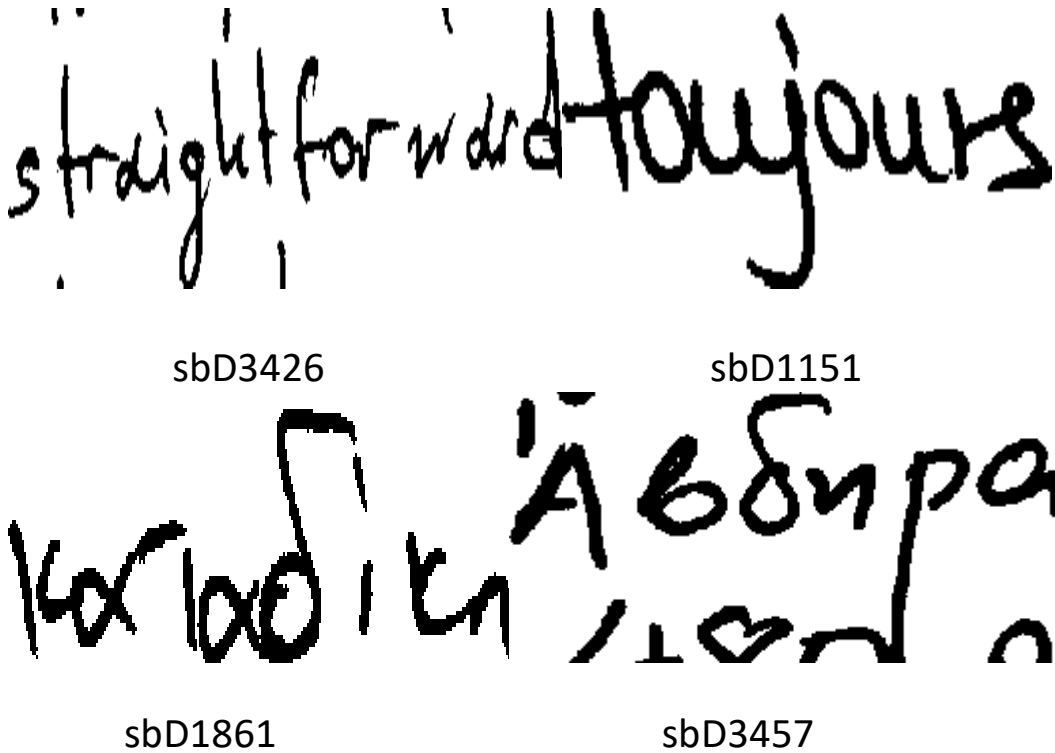


Figure 3.4| Échantillons des mots de requête sélectionnés à partir de la base de données moderne

Dans la première série d'expériences, nous étudions l'effet du paramètre d'échelle σ ($\sigma \in \{1,2,4,8,16\}$) dans le calcul des caractéristiques oBIFs. La distance de **City-block** a été utilisée comme métrique de distance pour de matching, dans ces expériences tandis que le paramètre ϵ a été défini sur une petite valeur **0,001** après-avoir essayé les valeurs **0,1** et **0,01** et nous avons constaté que le paramètre **0,001** donne de meilleurs résultats, comme indiqué dans le **Table 3.1.**

Le **Table 3.2** résume les performances en termes de **Top1, Top5, Top10** et **mAP.**

Paramètre oBIFs $\sigma = [8-16]$	Top 1	Top 5	Top 10	mAP
$\epsilon = 0.1$	12,88	18,58	20,38	23,39
$\epsilon = 0.01$	13,32	19,11	20,84	23,83
$\epsilon = 0.001$	14.06	19.85	21.7	24.72

Table 3.1| Comparaison entre que le paramètre $\epsilon = 0.1, \epsilon = 0.01$ et $\epsilon = 0.001$

Chapitr III : Méthode proposé et résultats expérimentaux

On peut voir dans le **Table 3.2** que les performances d'appariement varient en fonction du paramètre d'échelle dans le calcul des **oBIFs**. Les fonctionnalités **oBIFs** générées et combinées à l'aide des valeurs du paramètre d'échelle $\sigma = [8,16]$ surpassent les autres configurations rapportant une **mAP** de **24,72 %**. La performance sur d'autres mesures indique une précision **Top1** de **14.06 %** ,**Top5** de **19.85 %** et **Top10** de **21.7 %**.Lestaux de performance de ces expériences sont résumés dans la **Figure 3.4**.

ParamètreoBIFs		Performances (%) City-block				
ϵ	Σ	Top 1	Top 5	Top 10	mAP	
0.001	1	09.87	12.87	140.0	17.06	
	2	10.07	13.50	14.80	18.00	
	4	11.13	14.70	16.00	18.74	
	8	12.30	17.03	18.60	21.96	
	16	13.64	19.33	21.01	24.09	
	1 2	10.05	13.17	14.49	17.46	
	1 4	10.97	14.50	15.79	18.53	
	1 8	12.22	17.07	18.53	21.55	
	1 16	13.28	19.00	20.85	23.93	
	2 4	11.12	14.63	15.86	18.62	
	2 8	12.13	16.91	18.49	21.54	
	2 16	13.06	18.43	20.26	22.89	
	4 8	12.42	16.77	18.41	21.33	
	4 16	12.89	17.92	19.71	22.60	
	8 16		14.06	19.85	21.70	24.72
	1 2 4		11.01	14.56	15.82	18.60
	1 4 8		12.31	17.14	18.45	21.31
	1 4 16		12.73	17.75	19.58	22.44
	1 8 16		13.86	19.64	21.53	24.54
	2 4 8		12.42	16.70	18.35	21.19
	2 8 16		13.69	19.52	21.30	24.28
	4 8 16		13.65	19.02	20.76	23.63
	1 2 4 8		12.61	16.76	18.33	21.23
	1 4 8 16		13.61	18.97	20.66	23.54
	1 2 8 16		13.55	19.36	21.12	24.12
	2 4 8 16		16.19	18.70	20.37	23.26
	1 2 4 16		12.64	17.51	19.19	22.02
	1 2 4 8 16		13.35	18.68	20.35	23.16

Table 3.2| Performances avec différents paramètres dans le calcul des **oBIFs**

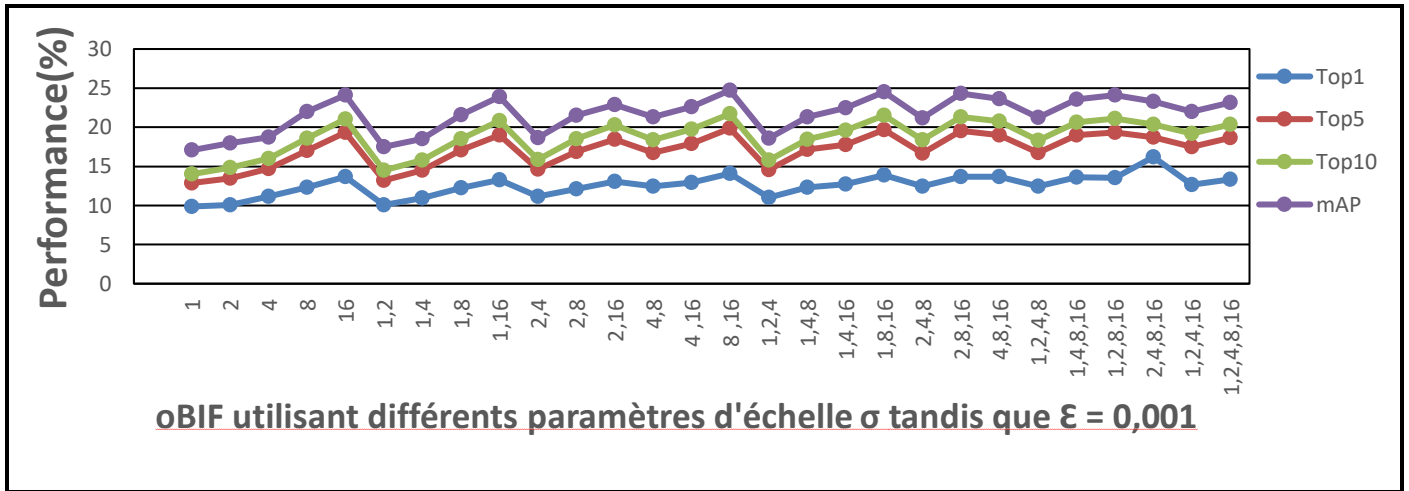


Figure3.5| Performance de la méthode proposée à l'aide des oBIFs

En plus de la distance de **City-block**, nous avons évalué la meilleure configuration d'oBIFs ($\sigma = [8-16]$) à l'aide de différentes distances, et les résultats sont présentés dans **le Table 3.3**.

Paramètre oBIFs	Type de distance	Top 1	Top 5	Top 10	mAP
$\epsilon=0.001$ Et $\sigma = [8-16]$	Correlation	13.17	18.98	20.85	23.92
	Euclidean	12.88	18.61	20.26	23.38
	Cosine	13.22	19.08	20.77	23.87
	Chebychev	11.33	15.82	17.50	20.42
	Spearman	14.16	19.76	21.67	24.71

Table 3.3| Performance de la méthode proposée en fonction de différentes métriques de distance

3.4. Comparaison entre les travaux de compétitions et la méthode proposée

Dans Cette comparaison des travaux de Competitionc2014 (H-KWS 2014) et la méthode qui nous proposée est présenté dans **le Table 3.5**. On peut voir le classement de la méthode proposée en classée quatrième .Cela ne signifie pas que la méthode proposée est inefficace.

Chapitr III : Méthode proposé et résultats expérimentaux

Dans les travaux de compétition (H-KWS 2014) utilise le prétraitement des images de mots, mais dans la méthode qui proposait, les images de mots ne sont pas prétraitées et les caractéristiques sont extraites directement des images brutes.

Par conséquent, les résultats encouragent et soutiennent l'utilisation des oBIFs pour caractériser les images de mots et je pense que dans le cas du prétraitement des images de mots, de bons résultats seront donnés et surpasseront les autres méthodes.

Méthodes	Extraction des caractéristique	Maching (Distance)	Basede données	mAP (%)
H-KWS 2014	SIFT	SVM	Modren	52.3
H-KWS 2014	LBP et HOG	Euclidien	Modren	33.8
H-KWS 2014	SIFT / SURF	Euclidien	Modren	27.8
Méthode proposée	oBIFs	City-block	Modren	24,72

Table 3.4 | Performance de comparaison entre les travaux de compétitions et la méthode proposée

Conclusion :

Ces résultats sont encourageants et soutiennent l'utilisation des oBIFs pour caractériser les images de mots. Il convient également de noter qu'aucun prétraitement n'est effectué sur les images de mots et que les caractéristiques sont extraites directement des images brutes.

CONCLUSION GÉNÉRALE

Conclusion Générale :

Le domaine de repérage des mots dans les documents manuscrits est un domaine de recherche exploré depuis plusieurs décennies et débouche sur un nombre important de travaux de recherche. Il y a une amélioration de résultats, mais jusqu'à présent aucune solution parfaite au problème de repérage des mots dans les documents manuscrits qui achève un taux de reconnaissance de 100 % d'un manière général. Pour cela le domaine de repérage des mots reste un sujet de recherche ouvert. L'objectif de ce travail est la réalisation d'un système de repérage des mots dans les documents manuscrits modernes.

Durant cette recherche, nous avons traversé aux deux étapes classiques d'un système de repérage des mots, dans la phase d'extraction des caractéristiques nous utilisons l'**oBIFs** et utiliser la distance **City-block** dans la phase matching, ce choix est justifié par la simplicité et l'efficacité de la méthode.

Afin de valider notre système, une étude expérimentale est réalisée à l'aide de la base de données de la compétition de repérage des mots **ICFHR 2014** et des résultats prometteurs sont rapportés en termes de précision **P@K** et de **meanAveragePrecision(mAP)**.

Cette étude a ciblé le problème du repérage des mots dans les documents manuscrits et a examiné l'efficacité des caractéristiques d'image de base orientées (**oBIFs**) pour caractériser les mots manuscrits.

Les premiers résultats obtenus sont encourageants et valident l'efficacité des **oBIFs** dans la caractérisation des images de mots. À noter également que nous ne réalisons aucun **pré-traitement** dans le mot images et que les caractéristiques sont directement extraites des images brutes.

Perspectives :

Malgré les résultats obtenus, un certain nombre d'amélioration de performance et des extensions peut être envisagé :

- ❖ Utiliser des techniques de prétraitements des images.
- ❖ Améliorer le system par utiliser d'autres méthodes de matching ou par combiner les matching avec d'autres ou plusieurs classifier.
- ❖ Utiliser d'autres méthode d'extraction des caractéristiques à la place oBIFs.
- ❖ Tester avec d'autres bases de données en introduisant un plus grand nombre des scripteurs, pour inclure plus de variations dans les styles d'écritures, ainsi la généralisation pourra être plus performante.

RÉFÉRENCE BIBLIOGRAPHIE

Références bibliographiques :

- [1] C. C. Tappert, C. Y. Suen, and T. Wakahara, "The state of the art in online handwriting recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 12, no. 8, pp. 787–808, Aug. 1990, doi: 10.1109/34.57669.
- [2] Kaur, A., Baghla, S., & Kumar, S. (2015). Study of various character segmentation techniques for handwritten off-line cursive words: A review. *International Journal of Advances in Science Engineering and Technology*, 3(3), 154-158.
- [3] Amin, A., Al-Sadoun, H., & Fischer, S. (1996). Hand-printed Arabic character recognition system using an artificial network. *Pattern recognition*, 29(4), 663-675.
- [4] Manmatha, R., Han, C., & Riseman, E. M. (1996, June). Word spotting: A new approach to indexing handwriting. In *Proceedings CVPR IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (pp. 631-637). IEEE.
- [5] Cao, H., & Govindaraju, V. (2007). Template-free word spotting in low-quality manuscripts. In *Advances In Pattern Recognition* (pp. 135-139).
- [6] Amit and Rishi, Rahul Choudhary and Savita Ahlawat, "A new character segmentation approach for off-line cursive handwritten words," *Procedia Computer Science*, vol. 17, pp. E88-95, 2013.
- [7] Chergui, L. (2013). Combinaison de classifieurs pour la reconnaissance de mots arabes manuscrits. *Université de Mentouri, Constantine*.
- [8] Kefali, A., Sari, T., & Sellami, M. (2009). Evaluation de plusieurs techniques de seuillage d'images de documents arabes anciens. In *5ème symposium international Images Multimédias Applications Graphiques et Environnements* (pp. 123-134).
- [9] ABA, S., & Maroua, C. H. I. K. H. (2016). *Intitulé Reconnaissance Des Mots Arabes Manuscrits* (Doctoral dissertation, Université laarbitebessitebessa).
- [10] Trier, O. D., & Taxt, T. (1995). Evaluation of binarization methods for document images. *IEEE transactions on pattern analysis and machine intelligence*, 17(3), 312-315.
- [11] introduisez à ce niveau le thème de votre projet De fin d'études Nature du document Réalisé par Encadré parPromotion : 2007 / 2008 2010/2011 BOUDRAA Omar OULAD KADDOUR Mohamed Mr SEHAD Abdnour Pro CHIBANI YUCEF Recherche d'informations dans les documents anciens historiques par le word spotting.

Références bibliographiques

- [12] Grandidier, F. (2003). *Un nouvel algorithme de sélection de caractéristiques: application à la lecture automatique de l'écriture manuscrite* (Doctoral dissertation, École de technologie supérieure).
- [13] Bow, S. T. (2002). *Pattern recognition and image preprocessing*. CRC press.
- [14] Belaïd, A. (2001). Reconnaissance automatique de l'écriture et du document. *Campus scientifique, Vandoeuvre-Lès-nancy*.
- [15] A New Character Segmentation Approach for Off-Line Cursive Handwritten Words
- [16] Choudhary, A., Rishi, R., & Ahlawat, S. (2013). A new character segmentation approach for off-line cursive handwritten words. *Procedia Computer Science, 17*, 88-95.
- [17] Filtre médian [RONSE, 2009]
- [18] <https://arthurflor23.medium.com/text-segmentation-b32503ef2613>
- [19] Antonacopoulos, et al., 2005
- [20] Ahonen, T., Hadid, A., & Pietikäinen, M. (2004, May). Face recognition with local binary patterns. In *European conference on computer vision* (pp. 469-481). Springer, Berlin, Heidelberg.
- [21] Dalal, N., & Triggs, B. (2005, June). Histograms of oriented gradients for human detection. In *2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05)* (Vol. 1, pp. 886-893). Ieee.
- [22] Dalal, N. (2006). *Finding people in images and videos* (Doctoral dissertation, Institut National Polytechnique de Grenoble-INPG).
- [23] Hossain, M. A., & Afrin, S. (2019). Optical Character Recognition based on Template Matching. *Global Journal of Computer Science and Technology*.
- [24] Fejer, M. M., Magel, G. A., Jundt, D. H., & Byer, R. L. (1992). Quasi-phase-matched second harmonic generation: tuning and tolerances. *IEEE Journal of quantum electronics, 28*(11), 2631-2654.
- [25] Székely, G. J., Rizzo, M. L., & Bakirov, N. K. (2007). Measuring and testing dependence by correlation of distances. *The annals of statistics, 35*(6), 2769-2794.
- [26] Singhal, Amit (2001). " Recherche d'informations moderne : un bref aperçu ". *Bulletin du comité technique de l'IEEE Computer Society sur l'ingénierie des données 24* (4) : 35-43.
- [27] Danielsson, P. E. (1980). Euclidean distance mapping. *Computer Graphics and image processing, 14*(3), 227-248.
- [28] Rothfeder, J. L., Feng, S., & Rath, T. M. (2003, June). Using corner feature correspondences to rank word images by similarity. In *2003 Conference on Computer Vision and Pattern Recognition Workshop* (Vol. 3, pp. 30-30). IEEE.

- [29] Lladós, J., Rusinol, M., Fornés, A., Fernández, D., & Dutta, A. (2012). On the influence of word representations for handwritten word spotting in historical documents. *International journal of pattern recognition and artificial intelligence*, 26(05), 1263002.
- [30] Aldavert, D., Rusinol, M., Toledo, R., & Lladós, J. (2013, August). Integrating visual and textual cues for query-by-string word spotting. In *2013 12th International conference on document analysis and recognition* (pp. 511-515). IEEE.
- [31] Almazán, J., & Gordo, A. (2013). Words with attributes library.
- [32] Kovalchuk, A., Wolf, L., & Dershowitz, N. (2014, September). A simple and fast word spotting method. In *2014 14th international conference on frontiers in handwriting recognition* (pp. 3-8). IEEE.
- [33] Yao, S., Wen, Y., & Lu, Y. (2015, August). Hog based two-directional dynamic time warping for handwritten word spotting. In *2015 13th International Conference on Document Analysis and Recognition (ICDAR)* (pp. 161-165). IEEE.
- [34] Retsinas, G., Louloudis, G., Stamatopoulos, N., & Gatos, B. (2018). Efficient learning-free keyword spotting. *IEEE transactions on pattern analysis and machine intelligence*, 41(7), 1587-1600.
- [35] Pratikakis, I., Zagoris, K., Gatos, B., Louloudis, G., & Stamatopoulos, N. (2014, September). ICFHR 2014 competition on handwritten keyword spotting (H-KWS 2014). In *2014 14th International Conference on Frontiers in Handwriting Recognition* (pp. 814-819). IEEE.
- [36] Gatos, B., Stamatopoulos, N., & Louloudis, G. (2011). ICDAR2009 handwriting segmentation contest. *International Journal on Document Analysis and Recognition (IJ DAR)*, 14(1), 25-33.
- [37] Newell, A. J., Griffin, L. D., Morgan, R. M., & Bull, P. A. (2010, December). Texture-based estimation of physical characteristics of sand grains. In *2010 International Conference on Digital Image Computing: Techniques and Applications* (pp. 504-509). IEEE.
- [38] Newell, A. J., & Griffin, L. D. (2014). Writer identification using oriented basic image features and the delta encoding. *Pattern Recognition*, 47(6), 2255-2265.
- [39] Gattal, A., Djeddi, C., Chibani, Y., & Siddiqi, I. (2016, April). Isolated handwritten digit recognition using oBIFs and background features. In *2016 12th IAPR Workshop on Document Analysis Systems (DAS)* (pp. 305-310). IEEE.

Références bibliographiques

- [40] Abdeljalil, G., Djeddi, C., Siddiqi, I., & Al-Maadeed, S. (2018, August). Writer identification on historical documents using oriented basic image features. In *2018 16th International Conference on Frontiers in Handwriting Recognition (ICFHR)* (pp. 369-373). IEEE.
- [41] Abbas, F., Gattal, A., Djeddi, C., Siddiqi, I., Bensefia, A., & Saoudi, K. Texture feature column scheme for single and multi-script writer identification. *IET Biometrics* 10 (2), 179–193 (2021).
- [42] Gattal, A., Djeddi, C., Siddiqi, I., & Chibani, Y. (2018). Gender classification from offline multi-script handwriting images using oriented basic image features (oBIFs). *Expert Systems with Applications*, 99, 155-167.
- [43] Gahmousse, A., Gattal, A., Djeddi, C., & Siddiqi, I. (2020, October). Handwriting based Personality Identification using Textural Features. In *2020 International Conference on Data Analytics for Business and Industry: Way Towards a Sustainable Economy (ICDABI)* (pp. 1-6). IEEE.
- [44] Yousfi, D., Gattal, A., Djeddi, C., Siddiqi, I., & Bensefia, A. (2022). Keyword Spotting in Modern Handwritten Documents Using oBIFs. In *Mediterranean Conference on Pattern Recognition and Artificial Intelligence* (pp. 240-250). Springer, Cham.
- [45] Mitra, D., Sarkar, P., & Roy, P. (2019). Face Recognition by City-Block Distance Classifier in Supervised Machine Learning. *IJRAR DOI: <http://doi.org/10.1729/Journal>, 21653.*
- [46] <http://www.mathworks.com>
- [47] L. Ott, *Initiation MATLAB 1, SCALab.*
- [48] D. Jean-Luc, *Introduction à MATLAB, Université de Picardie.*
- [49] Wang, P., Eglin, V., Garcia, C., Llargeron, C., Lladós, J., & Fornés, A. (2014, April). A novel learning-free word spotting approach based on graph representation. In *2014 11th IAPR International Workshop on Document Analysis Systems* (pp. 207-211). IEEE.