

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 de Mathématiques et d'Informatique

Mémoire de master

Domaine : Mathématiques et Informatique

Filière : Informatique

Option : Systèmes et Multimédia

Thème :

*Combinaison de classifieurs pour l'identification
des traits de personnalité*

Présenté par :

YOUSFI RABEB

Devant le jury :

<i>Mr. ZEGARI AHMED</i>	<i>Maitre de conférences B</i>	<i>Université de Tébessa</i>	<i>Président</i>
<i>Mr. HADJEDJ ISMAIL</i>	<i>Maitres de conférences B</i>	<i>Université de Tébessa</i>	<i>Examineur</i>
<i>Mr. CHAWKI DJEDDI</i>	<i>Maitres de conférences A</i>	<i>Université de Tébessa</i>	<i>Rapporteur</i>

Année Universitaire : 2020-2021

Résumé

Dans ce manuscrit, un système automatique pour l'identification des traits de personnalité à partir de documents manuscrits est introduit. Le système proposé comprend deux étapes principales: l'extraction de caractéristiques et la classification (identification de traits de personnalité). Dans la première étape, les distributions des charnières des contours en utilisant différentes longueurs de segments sont extraites à partir de documents manuscrits. Dans la deuxième étape, nous avons utilisé les forêts aléatoires, les machines à vecteurs de support ainsi qu'un algorithme de boosting adaptatif pour la classification. Les expérimentations sont menées sur un ensemble de données qui comprend 570 documents manuscrits. Le taux d'identification le plus élevé est obtenu par l'utilisation de l'algorithme de booster adaptatif.

Keywords : Identification de traits de personnalité, documents manuscrits, charnières des contours, forêts aléatoires, machines à vecteurs de support, algorithme de boosting adaptatif.

Abstract

In this manuscript, a personality traits identification system based on handwritten documents is introduced. The proposed system consists of two main stages: feature extraction and classification (personality traits identification). In the first step, edge-hinge features are extracted from handwritten documents. In the second step, we have used Random Forests, Support Vector Machines and Adaboost as classifiers. The experiments are conducted on a database that includes 570 handwritten documents. The highest identification rate is achieved by the use of Adaboost classifier.

Keywords: Personality traits identification, handwritten documents, edge-hinge, random forests, support vector machines, adaboost.

الملخص في هذا العمل، تم تقديم نظام للتعرف على ملامح الشخصية باستعمال وثائق مكتوبة بخط اليد. يتكون النظام المقترح من مرحلتين أساسيتين: استخراج المميزات وتصنيفها (التعرف على ملامح الشخصية). في الخطوة الأولى، يتم استخراج ملامح الحافة من الوثائق المكتوبة بخط اليد. خلال المرحلة الثانية، استخدمنا الغابات العشوائية، آلة المتجه الداعم وخوارزمية التعزيز التكييفي للتصنيف. أجريت التجارب على قاعدة بيانات تحتوي على 570 وثيقة مكتوبة بخط اليد. تم تحقيق أعلى دقة من خلال استخدام خوارزمية التعزيز التكييفي.

الكلمات المفتاحية: التعرف على ملامح الشخصية، وثائق مكتوبة بخط، مرفصلات الحافة، الغابات العشوائية، آلة المتجه الداعم، خوارزمية التعزيز التكييفي.

Remercîment

Tout d'abord, nous remercions Dieu Tout-Puissant pour notre patience, notre bien-être et notre volonté. Nous tenons à exprimer nos remerciements à notre encadreur Mr.” **DJEDDI CHAWKI**”. Nous tenons également à lui exprimer notre gratitude pour sa haute éthique et son soutien à notre égard afin que notre thèse porte ses fruits. Nous tenons également à remercier Mr . “**ABDULLATIF GAHMUS**” pour sa haute éthique et ses conseils et efforts afin que cette thèse puisse être menée à bien. Nous tenons également à remercier les enseignants du Département de mathématiques et d'informatique qui ont partagé leurs connaissances et leur expérience avec nous. Nous remercions également tous les membres du jury d'avoir accepté d'être présents à la présentation de cette thèse. Sans oublier de remercier nos parents pour tous les sacrifices qu'ils ont consentis pour nous.

Enfin, nous remercions sincèrement tous ceux qui nous ont toujours encouragés lors de la préparation de cette thèse.

Sommaire

1	Introduction Générale.....	1
1.1	Contexte du travail.....	1
1.2	Objectifs du travail.....	3
1.3	Organisation de la thèse.....	3
2	Analyse et classification automatique d'images de documents manuscrits.....	5
2.1	Introduction.....	5
2.2	Analyse graphologique.....	7
2.3	Modèle à cinq facteurs.....	9
2.4	Reconnaissance de scripteurs à partir de l'écriture manuscrite.....	15
2.5	Prédiction des troubles neurologiques à partir de l'écriture manuscrite.....	17
2.6	Reconnaissance des émotions à partir de l'écriture manuscrite	19
2.7	Prédiction de la personnalité à partir de l'écriture manuscrite.....	21
1.7.1	Travaux connexes.....	22
2.8	Analyse critique.....	29
2.9	Conclusion.....	29
3	Caractéristiques texturales pour l'identification de traits de personnalité à partir de documents manuscrits.....	30
3.1	Introduction.....	31
3.2	Ensemble des données.....	34
3.3	Extraction de caractéristiques.....	35
3.4	Classification.....	35
3.4.1	Les forêts aléatoires (RF).....	36
3.4.2	Les Machines à vecteurs de support (SVM).....	36
3.4.3	Algorithme de boosting adaptatif (ADABOOST).....	37
3.5	Résultats et discussion.....	37
3.6	Conclusion.....	41
4	Conclusion générale et perspectives.....	42
	Bibliographie.....	43

Liste de Figures

Figure 2.1 : Tableau d'analyse de l'écriture manuscrite 1.....	8
Figure 1.2 : Tableau d'analyse de l'écriture manuscrite2.....	9
Figure 3.2 : Représentation hiérarchique des traits de personnalité, y compris les facettes et sous-facettes.....	14
Figure 2.4 : une page de cahier d'un enfant de CE2	18
Figure 2.5 : Une page de cahier d'un enfant de CE2	19
Figure 2.6 : Indication d'émotion à partir des caractéristiques	20
Figure 3.1 : échantillons manuscrits.....	32
Figure 3.2 : Catégorisation des résultats FFM.....	33
Figure 3.3 : Description graphique des caractéristiques des charnières des contours.....	35

Liste de Tables

Tableau 3.1: Répartition des échantillons sur les cinq traits de personnalité.....	34
Tableau 3.2 : Une vue d'ensemble des caractéristiques implémentées et leurs dimensionnalités.....	35
Tableau3.3 : Taux globaux d'identification enregistrés en utilisant les forêts aléatoires (RF), les Machines à vecteurs de support (SVM) ainsi que l'algorithme de BOOSTING adaptatif –(ADABOOST).....	38
Tableau 3.4 : Résultats réalisés pour les cinq traits de personnalités en utilisant les forêts aléatoires (RF).....	39
Tableau 3.5 : Résultats réalisés pour les cinq traits de personnalités en utilisant les machines à vecteurs de support (SVM).....	39
Tableau 3.6 : Résultats réalisés pour les cinq traits de personnalités en utilisant l'algorithme de boosting adaptatif (ADABOOST).....	39
Tableau 3.7 : Comparaison des performances des systèmes de prédiction des traits de personnalité basés sur l'écriture manuscrite.....	40

INTRODUCTION

GENERALE

1.1. Contexte du travail

De nos jours, les gens communiquent entre eux via de multiples supports, mais l'écriture manuscrite est la plus ancienne méthode de communication non verbale utilisée depuis des siècles. L'écriture manuscrite varie d'une personne à l'autre et elle joue un rôle très important pour juger la personnalité de quelqu'un. L'écriture manuscrite est un processus continu, ce n'est pas quelque chose que les gens peuvent développer du jour au lendemain. Le style d'écriture d'une personne change selon l'humeur de cette personne. L'écriture manuscrite a tendance à être plus difficile à lire lorsqu'une personne est fatiguée, tandis que l'écriture manuscrite a tendance à être plus compréhensible si une personne est enthousiaste [1] [2].

Pratiquement 70 à 80 pour cent des cellules cérébrales travaillent activement pendant l'écriture alors que seulement 15 à 20 pour cent travaillent pendant la parole [3]. Lors de l'écriture, le signal sur ce qu'il faut écrire ou le contenu de l'écriture est donné par la partie consciente du cerveau. Le mouvement de la main pour produire l'écriture est contrôlé par des parties conscientes et inconscientes. Le comportement, les traits de personnalité, l'état physique, mental et émotionnel d'un écrivain peuvent être identifiés par l'écriture manuscrite. Un lien est créé entre le cerveau et l'écriture, ce qui se traduit par un mouvement neuromusculaire unique qui est lié au schéma du cerveau [4].

Cette raison fait que le terme écriture manuscrite est appelé «écriture cérébrale». L'écriture manuscrite en fait «l'écriture cérébrale» en tant qu'individu peut contrôler son cerveau concernant ce qu'il faut écrire en dehors de la façon d'écrire [5]. La façon dont une personne se comporte et

communiqué avec les autres est appelée personnalité. Chaque humain a une personnalité unique qui définit les aspects forts et faibles de cette personne. En connaissant la personnalité, nous pouvons en tirer de nombreux avantages. Par exemple, si nous connaissons la personnalité d'une personne, nous pouvons réduire la possibilité de conflits prolongés. La compréhension de la personnalité est essentielle pour déterminer l'interaction sociale d'une personne [6][7].

Il existe de nombreuses façons disponibles, comme le ton de la voix, l'expression du visage, le geste, etc., qui peuvent être utilisées pour déterminer la personnalité d'une personne. Pourtant, l'écriture manuscrite est une méthode encore plus puissante à cette fin. L'écriture manuscrite est un bras nécessaire pour déduire la personnalité d'un individu. Tout le monde ne peut pas facilement connaître la personnalité d'une personne grâce à l'écriture manuscrite, les personnes qui analysent l'écriture manuscrite pour identifier la personnalité d'un individu sont appelés graphologues.

La graphologie est considérée comme une forme innovante de psychologie. Depuis le début des années 1900, la personnalité de l'individu a attiré l'attention des psychologues [8]. Ils analysent différentes caractéristiques de l'écriture manuscrite telles que l'espace entre les lignes, la façon dont la lettre « t » ou « f » est écrite, la façon dont les lettres ou les mots sont écrits, la ligne de base et la taille des lettres, etc. pour évaluer un aspect psychologique d'un individu. Les caractéristiques de l'écriture manuscrite et leurs types ont été expliqués dans la section 2.5.3. Sur la base de règles graphologiques, le graphologue interprète les caractéristiques extraites [9],[10],[11]. Un humain peut avoir de nombreux types de caractéristiques telles que la maturité, la sociabilité, etc. Le graphologue essaie d'identifier ces caractéristiques en fonction des caractéristiques de l'écriture manuscrite. Par exemple, s'il y a un changement soudain de vitesse, de taille et de pression, alors cet individu sera sensible, alors que le scripteur mûrira s'il y a une répartition égale de l'espace entre les lignes et les mots. Le graphologue a une vaste portée dans différents domaines comme la psychologie, le diagnostic médical, le recrutement, etc.

Dans la plupart des panels de recrutement, les psychiatres font également intervenir et accèdent à la personnalité d'une personne interrogée. Si nous parlons de médecine, les graphologues jouent également un rôle essentiel ici, des maladies comme les troubles neurologiques peuvent être facilement prédites à des stades précoces en examinant l'écriture manuscrite. Différents aspects d'un individu peuvent être explorés par l'analyse de son écriture [12].

2. Objectifs du travail

Un certain nombre de travaux de recherche ont été menés à différentes institutions sur l'écriture manuscrite. Inspirés par la forte corrélation entre l'écriture manuscrite et la personnalité, nous avons l'intention de développer un système automatique pour l'identification de traits de personnalité d'un individu à partir d'images scannées de son écriture manuscrite. Le système proposé exploite les distributions des charnières des contours (EDGE-HINGE) comme caractéristiques. La classification est effectuée à l'aide des forêts aléatoires (RF), les Machines à vecteurs de support (SVM) ainsi qu'un algorithme de booster adaptatif (ADABOOST). Les expérimentations seront réalisées sur une base de données contenant 570 documents manuscrits.

3. Organisation du mémoire

Ce mémoire est structuré en deux parties. La première est consacrée à la présentation des principaux concepts, outils et travaux relatifs à l'étude entreprise. Dans la deuxième partie du mémoire, nous abordons de manière détaillée nos choix conceptuels, la mise en œuvre ainsi que les résultats obtenus par le système proposé pour l'identification des traits de personnalité à partir de documents manuscrits.

Chapitre 2. *Analyse et classification automatique d'images de documents manuscrits*

Ce chapitre est consacré à la présentation des quelques domaines de recherche liés directement avec le problème étudié. Il présente les travaux connexes dans le domaine de l'identification des traits de personnalité à partir d'images de

textes manuscrits. Nous décrivons les cinq grands traits de personnalité considérés dans le cadre de ce mémoire ainsi qu'une étude détaillée sur caractéristiques de l'écriture manuscrite et la personnalité. Nous terminons ce chapitre la présentation du processus d'analyse de l'écriture manuscrite pour détecter la personnalité d'une personne en se basant sur des documents manuscrits.

Chapitre 3. Caractéristiques texturales pour l'identification de traits de personnalité à partir de documents manuscrits

Ce chapitre se détache des aspects théoriques abordés dans le premier chapitre et s'oriente vers la présentation de notre contribution qui consiste en une approche indépendante du texte pour l'identification des traits de personnalité à partir de documents manuscrits en se basant sur la caractérisation des différentes images de textes par les distributions des charnières des contours (EDGE-HINGE). Nous décrivons la base de données utilisée avant de nous focaliser sur la présentation de la méthode d'extraction de caractéristiques proposée, les forêts aléatoires (RF), les machines à vecteurs de support (SVM) ainsi qu'un algorithme de booster adaptatif (ADABOOST) sont employés pour la classification. Les expérimentations effectuées seront aussi présentées. A la fin de ce chapitre, les résultats sont exposés et discutés.

A la fin de ce mémoire, nous émettons nos conclusions sur le travail que nous avons entrepris dans le domaine de l'identification des traits de personnalité à partir de documents manuscrits. Nous présentons aussi les perspectives d'extensions futures du travail que nous avons présenté dans ce document.

CHAPITRE

2

Analyse et classification automatique d'images de documents manuscrits

Ce chapitre est consacré à la présentation des quelques domaines de recherche liés directement avec le problème étudié. Il présente les travaux connexes dans le domaine de l'identification des traits de personnalité à partir d'images de textes manuscrits. Nous décrivons les cinq grands traits de personnalité considérés dans le cadre de ce mémoire ainsi qu'une étude détaillée sur les caractéristiques de l'écriture manuscrite et la personnalité. Nous terminons ce chapitre la présentation du processus d'analyse de l'écriture manuscrite pour détecter la personnalité d'une personne en se basant sur des documents manuscrits.

2.1. Introduction

La reconnaissance de la personnalité humaine devient de plus en plus importante dans le monde moderne. Cela aide les êtres humains à simplifier leur travail et à résoudre des problèmes plus complexes. Un être humain peut avoir de nombreuses caractéristiques uniques et l'écriture manuscrite prend des informations précieuses pour mieux comprendre les différents états de l'écrivain. On a récemment découvert que l'écriture manuscrite avait une forte corrélation avec le fonctionnement du cerveau et le côté psychologique des êtres humains. L'écriture manuscrite également appelée écriture cérébrale est une mesure utile pour identifier les traits de personnalité caractéristiques d'un individu et elle montre la relation entre l'écriture manuscrite en neuromusculaire et les effets de divers facteurs tels que le vieillissement et les problèmes de santé. L'analyse de l'écriture manuscrite ou de la graphologie est un processus

scientifique d'identification, de détermination et de compréhension de la personnalité à travers les traits et les modèles révélés par l'écriture.[13]

Les analystes de l'écriture, également appelés graphologues, peuvent examiner l'écriture d'un individu pour prédire les traits de personnalité de l'écrivain. En graphologie, l'écriture manuscrite est analysée par des éléments graphiques structurels afin d'en tirer des informations sur la personnalité de l'écrivain. À l'aide de la théorie de la graphologie, l'écriture manuscrite révèle la vraie personnalité, y compris les dépenses émotionnelles, les peurs, l'honnêteté, les défenses et bien d'autres [13]. Un grand modèle à cinq facteurs est considéré pour découvrir les différences individuelles. Les cinq grands traits de personnalité sont cinq grands domaines ou dimensions de la personnalité qui sont utilisés pour décrire la personnalité humaine [12]. Les graphologues identifient les qualités, traits, attitudes, sentiments ou postures qui semblent indiqués dans l'écriture manuscrite, ils cherchent en outre à comprendre comment ces aspects de l'individualité peuvent s'intégrer ensemble pour constituer l'organisation dynamique que nous reconnaissons comme la personnalité de cet écrivain. [13]

Automatiser l'identification des traits de personnalité d'un écrivain à partir de son écriture est aujourd'hui extrêmement essentielle en raison de l'immense développement de la technologie et de ses applications dans de nombreux domaines. L'identification des traits de personnalité d'un écrivain à partir de ses documents écrits s'applique dans de nombreux domaines, tels que le recrutement et la sélection du personnel, la détection des problèmes psychologiques, la détection des talents et des problèmes cachés [12].

L'automatisation de l'identification des traits de personnalité par l'écriture manuscrite se heurte à un certain nombre de sous-problèmes, comme la conception d'algorithmes pour analyser l'écriture manuscrite de différents individus, identifier les caractéristiques les plus pertinentes permettant de caractériser les différents styles d'écriture et afin de choisir l'automatique technique de classifications les mieux adaptées à ce problème [12].

2.2. Analyse graphologique

Graphologie, inférence de caractère à partir de l'écriture manuscrite d'une personne. La théorie derrière la graphologie est que l'écriture manuscrite est une expression de la personnalité, par conséquent, une analyse systématique de la façon dont les mots et les lettres sont formées peut révéler des traits de personnalité. Les graphologues notent ces éléments, il y a environ 300 caractéristiques comme la taille des lettres individuelles et le degré et la régularité de l'inclinaison, la zone supérieure ou la casse (comme dans l, t, h, etc.), la zone inférieure (comme dans g, y, p, etc.), espacement des mots et des lignes et ligne ondulée[14]. D'autres considérations de base sont l'aspect général et l'impression de l'écriture, la pression des traits vers le haut et vers le bas et la douceur de l'écriture. La compétence d'interprétation du graphologue réside dans l'art psychologique de comprendre le mélange particulier de caractéristiques de l'écriture manuscrite un expert est capable de voir l'écrivain « sortir de la page».

Le graphologue actuel est un professionnel dont la formation et l'éthique dépassent l'image que certains se font de celui qui « prédit » à travers une loupe suspendue au cou. Il est démontré qu'il existe une relation entre la personnalité, les conduites, le niveau intellectuel et le niveau volitif, le tempérament et le caractère dans l'écriture manuscrite [14].

Les graphologues professionnels opèrent, selon un code d'éthique strict et ces experts sont constamment en demande, ceux qui l'utilisent reconnaissent sa valeur sur le lieu de travail en tant que méthode supplémentaire de compréhension du caractère. C'est donc un outil extrêmement utile pour identifier la qualité et la capacité des talents et du potentiel d'un individu, notamment dans l'orientation professionnelle et l'amélioration des relations [15]. Documents d'autres modèles comportementaux ou intuitifs puissants, il n'est pas facile d'expliquer comment et pourquoi la graphologie fonctionne, néanmoins, elle continue d'être utilisée, respectée et appréciée par beaucoup, car elle permet d'obtenir un niveau de résultats élevé [16] .

Tableau d'analyse de l'écriture manuscrite : ces tableaux de graphologie sur les caractéristiques, les aspects et les traits de l'écriture manuscrite vous aideront tout au long du processus d'analyse.

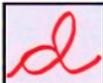
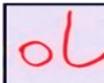
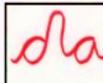
Traits	Exemples		
Lettre i	Forme de cercle	Coupant le l	Juste au-dessus de la tige
			
Lettre t	Croisé bas et court	Longues croix	En plein dans le mille
			
Lettre d	Grande boucle	Ovale séparé	D avec arcade
			
Lettre g	Boucle triangulaire	Boucle gonflée	Course verticale unique
			

Figure 2.1 : Tableau d'analyse de l'écriture manuscrite 1 [53]

L'analyse de l'écriture, selon la forme, la pression, la taille, l'inclinaison, les zones (zones supérieures, moyenne et inférieure), les barres en t, la disposition (marges et espaces entre les lettres, les mots et les lignes).

Une note manuscrite peut révéler des traits de personnalité cachés d'une personne, ce qui peut l'aider à connaître ses dépenses personnelles de base, ses faiblesses à améliorer sur une force sur laquelle s'appuyer. Voici quelques-uns de ces traits :

Caractéristiques	Traits d'écriture
Progressivité	Silencieux à droite, barres i dot et t placées haut, vitesse, lignes ascendantes.
Rationalité	Simplification, lignes droites, écriture verticale, barres en T fermes, points i placés bas.
Réceptivité	Haute pression avec une écriture irrégulière.
Réserve	Arcades, marge droite large, écriture verticale, marge inférieure large, écriture étroite avec pression.
Sensibilité	Inclinaison extrême droite, lignes vacillantes, pression avec irrégularité.
Sensualité	Boucles remplies d'encre, plénitude dans la zone inférieure.
La timidité	Absence d'ornementation, zones supérieures et inférieures surdimensionnées, large marge droite.
Sociabilité	Écriture connectée, guirlandes, oblique à droite, écriture large avec pression.
Spiritualité	Netteté, extensions exagérées dans la zone supérieure, légère pression.

Figure 2.2 : Tableau d'analyse de l'écriture manuscrite 2 [53]

2.3. Modèle à cinq facteurs:

Le modèle Big Five, également répandu sous le nom de modèle à cinq facteurs, est la conception de la personnalité la plus largement acceptée parmi les psychologues aujourd'hui. La théorie affirme que la personnalité peut être ramenée à cinq facteurs fondamentaux, connus sous l'acronyme CANOE ou OCEAN : les cinq grands traits de personnalité décrits par la théorie sont l'extraversion (également souvent orthographiée extroversion), l'agrément, l'ouverture, la conscience et le névrosisme [17] .

Ce modèle est bénéfique pour déterminer la personnalité d'une personne. Il s'agit d'une méthode basée sur un questionnaire, un type différent de questions

liées à la personnalité posées dans le questionnaire [18]. En remplissant le questionnaire FFM, la personne peut connaître la personnalité [19] .

Il est important de noter que l'un des cinq traits de personnalité représente la distance entre les deux extrêmes. Par exemple, l'extraversion représente un continuum entre l'extraversion extrémiste et l'introversion extrémiste. Par exemple, l'extraversion représente un continuum entre l'extraversion extrémiste et l'introversion extrémiste. Dans le monde réel, la plupart des gens sont entre deux pôles de chaque perspective [19] [20].

Les cinq facteurs sont issus de décennies de recherche sur la personnalité, proviennent de l'étude statistique des réponses aux items de la personnalité. En utilisant une technique appelée analyse factorielle, les chercheurs peuvent examiner les réponses des gens à des centaines d'items de personnalité et poser la question « quel est le mieux pour résumer un individu ? ». Cela a été fait avec de nombreux échantillons du monde entier et le résultat général est que, bien qu'il semble y avoir des variables de personnalité illimitées, cinq se démarquent du peloton en termes d'explication de nombreuses réponses de personnes à des questions sur leur personnalité, grandissant à partir des fondements des 16 facteurs de Cattell et devenant finalement le modèle de personnalité le plus accepté à ce jour. Ce modèle a été traduit en plusieurs langues et appliqué dans des dizaines de cultures, aboutissant à une recherche qui non seulement confirme sa validité en tant que théorie de la personnalité mais établit également sa validité au niveau international [20].

Ces cinq facteurs ne fournissent pas d'explications complètement exhaustives de la personnalité, mais ils sont connus sous le nom de Big Five car ils englobent une grande partie des termes liés à la personnalité. Les cinq facteurs ne sont pas nécessairement des traits en eux-mêmes, mais des facteurs dans lesquels s'inscrivent de nombreux traits et caractéristiques connexes [20] .

Par exemple, le facteur d'agrément englobe des termes comme la générosité, l'amabilité et la chaleur du côté positif et l'agressivité et l'humeur du côté

négatif. Tous ces traits et caractéristiques (et bien d'autres) constituent le facteur plus large de l'agrément [19].

Ci-dessous, nous expliquerons chaque facteur plus en détail et fournirons des exemples et des termes associés pour vous aider à avoir une idée des aspects et des particularités de la personnalité que ces facteurs couvrent.

Un acronyme populaire pour les Big Five est OCEAN. Les cinq facteurs sont présentés dans cet ordre ici [18][19][21].

✚ **Ouverture à l'expérience** (comprend des aspects tels que la curiosité intellectuelle et l'imagination créative)

A été décrit comme la profondeur et la complexité de la vie et des expériences mentales d'un individu [22]. On l'appelle aussi parfois intellect ou imagination.

L'ouverture à l'expérience concerne la volonté des gens d'essayer de nouvelles choses, leur capacité à être vulnérables et leur capacité à sortir des sentiers battus. Une personne très ouverte à l'expérience est probablement une personne qui aime apprendre, aime les arts, s'engage dans une carrière créative ou un passe-temps et aime rencontrer de nouvelles personnes [23].

✚ **Conscienciosité** (organisation, productivité, responsabilité)

Est un trait qui peut être décrit comme la tendance à contrôler les impulsions et à agir de manière socialement acceptable, des comportements qui facilitent le comportement orienté vers un but.

Les personnes consciencieuses excellent dans leur capacité à retarder la gratification, à respecter les règles, à planifier et à s'organiser efficacement [22].

✚ **Extraversion** (sociabilité, assertif, son contraire est Introversion)

Ce facteur a deux extrémités familières de son spectre : l'extraversion et l'introversion.

Il s'agit de savoir d'où un individu tire son énergie et comment il interagit avec les autres. En général, les extravertis puisent leur énergie ou se rechargent en interagissant avec les autres, tandis que les introvertis se fatiguent d'interagir avec les autres et reconstituent leur énergie avec la solitude [22].

✚ **Agréabilité** (compassion, respect, confiance dans les autres) :

Ce facteur concerne la manière dont les gens s'entendent bien avec les autres. Alors que l'extraversion concerne les sources d'énergie et la poursuite des interactions avec les autres, l'agréabilité concerne l'orientation de chacun vers les autres. C'est une construction qui repose sur la manière dont un individu interagit généralement avec les autres [22].

✚ **Névrosisme** (tendances à l'anxiété et à la dépression) :

N'est pas un facteur de méchanceté ou d'incompétence, mais un facteur de confiance et d'être à l'aise dans sa peau. Cela englobe la stabilité émotionnelle et le tempérament général [22].

Bien que la simplicité du modèle cinq grands traits de personnalité soit l'une de ses caractéristiques les plus attrayantes, certains chercheurs ont fait valoir que la mesure de facteurs complexes tels que la « performance au travail » nécessite un modèle plus détaillé de la personnalité humaine. À la lumière de cela, les universitaires ont cherché à améliorer le modèle ces dernières années.

En 2013, par exemple, les chercheurs Timothy Judge, JESSICA RODELL, Ryan Klinger, Lauren Simon et EEAN Crawford ont publié une version mise à jour du modèle qui décompose les cinq traits de personnalité en caractéristiques plus détaillées. Ils ont étiqueté ces « facettes » et « sous-facettes » [24] (Figure ci-dessous).

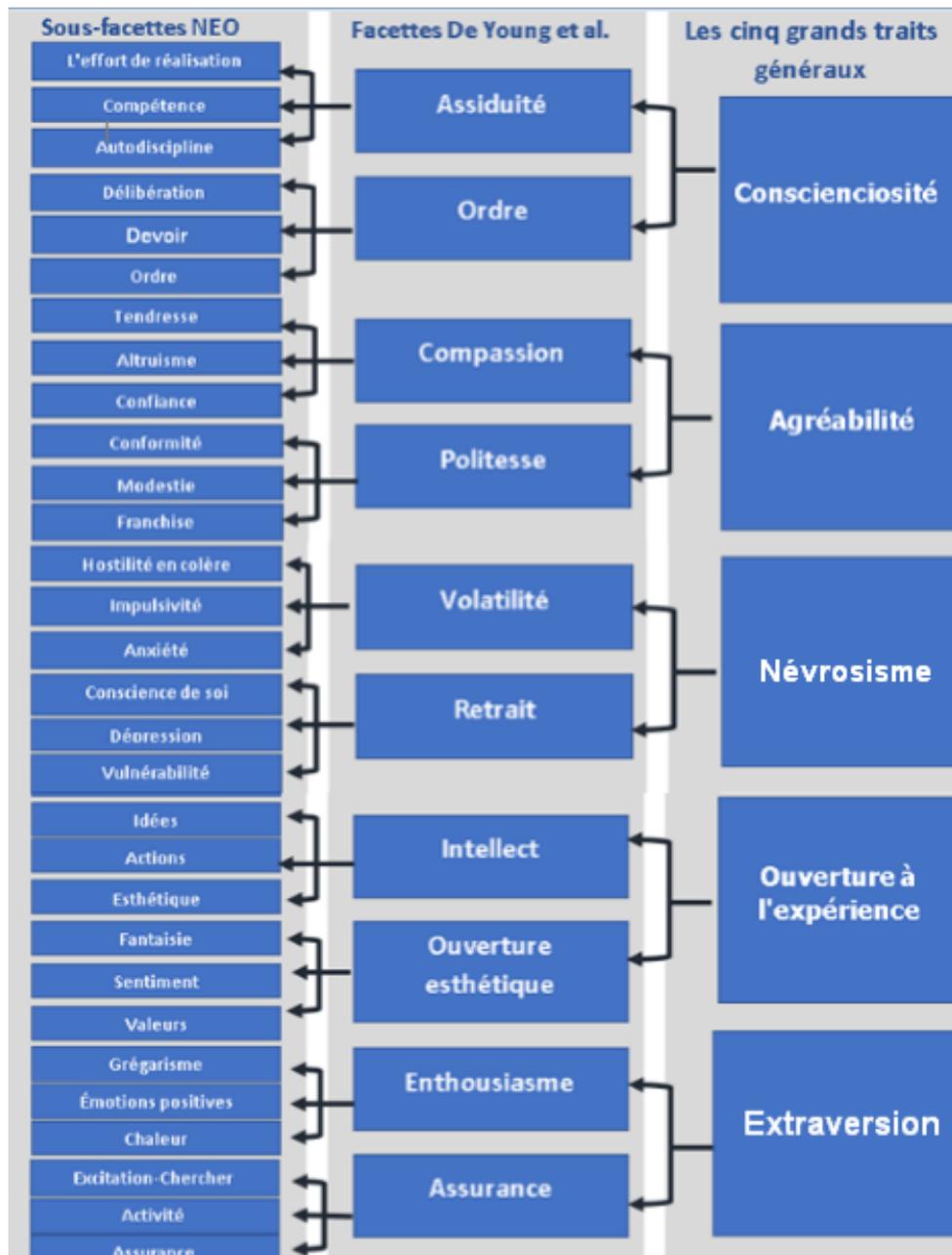


Figure 2.3 : Représentation hiérarchique des traits de personnalité, y compris les facettes et sous-facettes [54]

Comme le montre la figure 2.3, chaque trait de personnalité est divisé en deux facettes. Ces 10 facettes ont été proposées en 2007 par les psychologues Colin DEYOUNG, Lena QUILTY et Jordan Peterson dans leur article « Entre facettes et domaines: 10 aspects des cinq grands ». Cette version du modèle est connue sous le nom de Test de personnalité des cinq grands aspects.

L'équipe de chercheurs de Judge s'est appuyée sur le modèle à 10 facettes, en décomposant chaque facette encore plus en 30 sous-facettes (voir figure 2.3), qu'ils ont publié en 2013. Ces sous-facettes sont basées sur un modèle NEO PERSONALITY INVENTORY ©, qui a été développé pour la première fois en 1978 et révisé en 1990 par les psychologues Paul Costa et Robert McRae.

2.4. Reconnaissance de scripteurs à partir de l'écriture manuscrite

L'intérêt porté à l'écriture est grandissant, du fait de sa nature artistique intéressante, mais également du fait de la variété d'informations qu'elle peut véhiculer. En effet, certains s'attachent à faire parler les écritures pour qu'elles révèlent les traits comportementaux de leurs auteurs ou simplement leurs identités.

L'écriture manuscrite est considérée comme un acte personnel : chaque scripteur est caractérisé par son écriture, par la reproduction de détails et d'habitudes inconscientes. C'est pourquoi dans certains cas d'expertise l'analyse des échantillons d'écriture a même valeur que l'analyse des empreintes digitales. En effet, que l'authentification d'une personne à partir de son écriture manuscrite représente un domaine important de la biométrie, surtout lorsque l'écriture représente l'unique trace exploitable laissée par les personnes.

L'écriture désigne tout media conventionnel utilisé par des groupes d'individus pour communiquer entre eux. C'est une représentation du langage humain sous une forme graphique. Bien qu'elle puisse différer d'une population à une autre, ses concepts de base restent pratiquement les mêmes (alphabet, mots, phrases). L'écriture est facilement quantifiable, son acquisition peut se faire soit en mode dynamique (enligne) à l'aide d'un stylo électronique et d'une tablette d'acquisition numérique, soit en mode statique ou différé (hors-ligne) où son acquisition se fait au moyen d'un scanner ou d'une caméra [25] [26].

Rappelons que la vérification de l'écrivain consiste à déterminer si l'auteur identifié est celui qui a écrit. Identifier un écrivain consiste à faire correspondre

une entrée donnée à une entrée de référence dont nous connaissons l'auteur. Cette association se fait selon un indice de similarité.

Ces travaux postulent que c'est une caractéristique comportementale d'auto-expression, agissant sous l'influence de plusieurs facteurs propres, éducationnels, culturels, intentionnels, émotionnels, circonstanciels et contextuels. Ces facteurs sont le fruit d'un ensemble de propriétés :

- Le degré de maturité graphique de l'auteur (enfant, adulte, ...)
- La force d'expression graphique
- L'état psychologique du scripteur (fatigue, stress, ...)
- Le système d'écriture appris
- Les connaissances linguistiques
- Le matériel utilisé pour écrire (papier, encre, surface, ...)
- Le développement musculaire du scripteur

Si l'on considère une base composée de m échantillons d'écriture provenant de n écrivains différents, l'identification est définie par le fait qu'un échantillon d'écriture x inconnu appartient ou non à l'une des classes définies dans à partir des m échantillons d'écriture des n écrivains de cette base de données et si une correspondance a été trouvée, il est possible d'indiquer le nom de cet écrivain [26], [27].

Si l'identification de l'écrivain fait partie du même problème que la reconnaissance de l'écriture manuscrite, elle ne semble pas poser le même genre de difficulté. En effet, la tâche d'identification peut profiter de la variabilité des écrits pour les discriminer, tandis que la tâche de reconnaissance doit, au contraire, parvenir à surmonter la variabilité entre écrivains pour identifier le message textuel quel que soit l'auteur. La variabilité de l'écriture manuscrite d'un écrivain dépend de multiples facteurs psychologiques difficiles à déterminer avec précision [28].

2.5. Prédiction des troubles neurologiques à partir de l'écriture manuscrite

La dégradation de l'écriture manuscrite des personnes est principalement due à des pathologies telles que la maladie d'Alzheimer, la maladie de Parkinson, le cancer, les maladies cardiaques et d'éventuelles pathologies psychiatriques telles que la schizophrénie [29] .

Pour les enfants atteints de dyspraxie visuo-spatiale, c'est-à-dire de dyspraxie associée à un trouble neuro-visuel et spatial, c'est la difficulté spatiale d'organisation des lignes et des boucles qui met les enfants en difficulté. Les déformations de leur écriture ne sont donc pas stables, ce qui rend la relecture difficile, voire impossible, y compris par l'enfant lui-même [30].

Un enfant dyspraxique n'automatise pas la calligraphie, il est donc en permanence sous contrôle volontaire et donc coûteux en soins. L'enfant a alors une double tâche, il est possible d'effectuer deux tâches distinctes en même temps tant qu'au moins une des deux tâches est automatisée. Si aucun des deux ne l'est, les faire ensemble entraîne une baisse significative des performances ou un échec dans chacune des deux tâches. Par exemple, si je vous demande de marcher sur une poutre (ce qui n'est pas un processus automatisé pour tout le monde), vous allez pouvoir le faire si je ne vous demande rien d'autre. Si je vous demande de faire de même tout en effectuant un calcul en même temps esprit, vous ferez l'un ou l'autre. Ou bien vous tomberez de la poutre ! [31].

L'enfant dyspraxique automatisant mal ou pas du tout la calligraphie, il devra donc automatiser une mesure palliative. Le clavier sera donc proposé, dont la saisie devra être automatisée pour ne pas être dupliquée. Si votre saisie n'est pas automatisée, vous êtes exactement dans la même situation que si vous deviez dessiner les lettres à la main. Voici des lettres tirées d'un texte (lettres soulignées Figure 2.4). Ils semblent avoir globalement la même forme.

Supprimer du texte dans son ensemble, il est difficile de savoir de quelles lettres il s'agit. Mais on se dit qu'une fois qu'on a repéré cette lettre dans le texte, on peut la nommer [32].

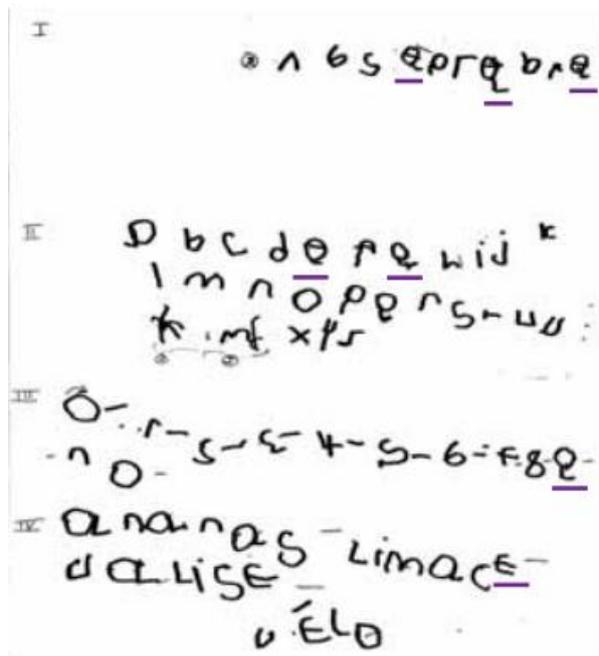


Figure 2.4 : Une page de cahier d'un enfant de CE2 [32]

Dans un article de Claire LE LOSTEC (Ergothérapeute, L'ADAPT, Paris) nommé « Troubles graphiques et troubles de l'écriture chez les enfants et troubles de l'écriture chez les enfants dyspraxie et dyspraxie »

Là où il a mené une expérience avec un enfant souffrant de troubles de l'écriture et là, il a expliqué ce qui suit : [32]

« Quentin, 6 ans et demi (Figure 2.5 page suivante). Quand on lui demande d'épeler son prénom spontanément, on a le « q » de Quentin, le « u » n'est pas fini, et il ne peut pas faire le dernier « n » à la fin. On lui demande alors d'écrire « maman », qu'il peut épeler parfaitement.

C'est ainsi qu'il organise sa « maman » de droite à gauche. Il s'en rend compte, alors il s'arrête. On lui propose alors d'écrire « papa » : le « p » est correct, comme le « q » de Quentin, mais il ne peut pas faire le 2e. En revanche, il est organisé dans le bon sens. En fait, il ne peut que se cibler sur un paramètre : s'il s'attache à aller dans la bonne direction, son graphisme se dégrade. Par contre, il est très critique de sa production : il dit que c'est faux, que ce n'est

pas joli, etc. Que les enfants puissent juger de leur production est une partie importante du diagnostic. Entre ce qu'ils ont imaginé pouvoir faire et le résultat sur la feuille, ils sont capables de dire que ce n'est pas du tout ce qu'ils voulaient faire. Ils sont capables de vous décrire très précisément ce qu'ils auraient aimé dessiner ou écrire, lorsque le résultat n'est pas concluant. » [32].



Figure 2.5 : Une page de cahier d'un enfant de CE2 [32]

2.6. Reconnaissance des émotions à partir de l'écriture manuscrite

L'analyse de l'écriture manuscrite est la technique utilisée pour mieux comprendre une personne grâce à son écriture. En examinant l'écriture, nous pouvons développer un croquis qui reflète les dépenses émotionnelles de l'écrivain, ses peurs, son honnêteté, son état mental et de nombreux autres traits de personnalité. Les émotions comprennent l'interprétation, la perception et la réponse des sentiments liés à l'expérience d'une situation particulière. Ce sont eux qui relient les pensées, les sentiments et les actions. Analysez les caractéristiques de l'écriture manuscrite telles que la ligne de base, l'inclinaison, la pression du stylo, la taille, la marge et la zone pour déterminer les niveaux d'émotion d'une personne. Cela aidera les personnes qui sont émotionnellement perturbées ou déprimées et qui ont besoin d'une aide psychologique pour surmonter ces émotions négatives [33] [34].

L'écriture manuscrite agit comme une fenêtre ouverte à travers laquelle toute la vision de l'esprit intérieur de l'écrivain peut être observée. Les émotions sont également associées à l'humeur, au tempérament, à la personnalité, aux perspectives et à la motivation [35].

Les émotions sont manifestement révélées par l'écriture manuscrite parce que l'écriture manuscrite est l'externalisation écrite de ce qui se passe en vous. Cela rend l'écriture manuscrite authentique qui ne peut pas être falsifiée. Une personne se sentant tendue, anxieuse ou déprimée représentera des traits tremblants dans certaines lettres, contrairement à la personne qui se sent heureuse et excitée aura des traits clairs, décorés et lisses. Une variation soudaine de l'écriture d'une personne peut être un signe d'ivresse de l'écrivain. De cette manière, l'analyse de l'écriture manuscrite est censée avoir la sensibilité d'un baromètre émotionnel et donne clairement une idée de l'excitation, de la peur, de l'anxiété, de l'irritabilité, de la dépression, de la colère et d'autres émotions similaires chez l'écrivain [33].

Caractéristique	Type	Émotion indiquée	Échantillon
Malformé Zone médiane	Milieu particulier Zone	Malheur, insatisfaction, anxiété, stress	
Malformé Zone supérieure	Déformé Boucles supérieures	Maladie mentale, psychose	
Tige pointue Boucles	Tige pointue Boucles	Colère, tension, anxiété	

Figure 2.6 : Indication d'émotion à partir des caractéristiques [56]

L'écriture manuscrite est unique à chaque personne et c'est une photographie instantanée de la façon dont nous pensons, ressentons et nous comportons à ce moment-là. C'est une forme de moment expressif figé sur le papier, à

travers lequel se révèlent des fonctions physiologiques et psychologiques. L'écriture d'une personne subit certaines variations avec son état d'esprit et ses émotions changeants. La disposition de la personne, qu'elle soit heureuse ou satisfaite, heureuse ou excitée, triste ou déçue, est exprimée à travers son écriture. Ces émotions peuvent être caractérisées par divers indicateurs d'écriture manuscrite ou des traits d'écriture tels que la ligne de base, l'inclinaison, la pression du stylet, la taille, les traits, l'espacement, les marges, les boucles, les points « i », la barre « t », etc. [36].

2.7. Prédiction de la personnalité à partir de l'écriture manuscrite

Le mode de vie des êtres humains a changé depuis l'ère numérique où tout peut être manipulé du bout du doigt, mais tous ces luxes pourraient revenir à une valeur de sécurité ou de fraude où masquer son identité avec un faux pas possible lors d'un cas avec écriture manuscrite. L'écriture manuscrite est exclusive à chaque personne comme une empreinte digitale est exclusive à chaque personne. Quelqu'un peut imiter l'écriture d'une autre personne pendant moins de quelques mots, ce qui la rend distinctive. L'écriture raconte le caractère de la personne, car l'écriture est couplée au cerveau et laisse inconsciemment un chemin concernant l'attribut de tempérament comme optimiste, pessimiste, équilibré, timide, etc..., qui pourrait être détecté. Différentes formes de styles d'écriture manuscrite sont prises en compte comme la pente ou l'angle de la phrase, le nombre de mots dans une région, l'inclinaison gauche ou droite de la phrase, etc [37] .

Comme avancé dans les sections précédentes, l'écriture manuscrite est unique pour chaque individu. Elle doit être identique et unique pour un scripteur, quel que soit le moyen utilisé pour la production de l'écriture. L'écriture manuscrite est écrite par le cerveau, pas par la main. Ainsi, l'écriture manuscrite est également appelée comme écriture cérébrale. Chaque trait de personnalité a un schéma neurologique dans le cerveau humain. Chaque schéma neurologique de

cerveau génère un mouvement neuromusculaire unique en son genre, identique pour chaque individu présentant ce trait de personnalité spécifique. Chaque mouvement en écriture révèle un trait de personnalité particulier. La graphologie est la science qui identifie ces traits dans l'écriture manuscrite et décrivant le trait de personnalité correspondant [38] .

Auparavant, l'analyse de l'écriture manuscrite se faisait manuellement en consacrant beaucoup de temps à prédire la nature de la personne. En analyse manuelle, la précision de l'analyse dépend des compétences du graphologue. Le graphologue est également enclin à la fatigue lorsque plusieurs échantillons doivent être analysés. Pour obtenir les services d'un graphologue expérimenté, les coûts sont élevés. D'autre part, l'analyse automatisée de l'écriture manuscrite est très rapide, précise, très peu coûteuse et pratique dans la prédiction de la personnalité humaine. [38]

Dans les sous sections suivantes, nous nous concentrons principalement sur les travaux connexes dans le domaine de l'identification des traits de personnalité à partir d'images de textes manuscrits, ensuite nous expliquons les caractéristiques de l'écriture manuscrite et la personnalité et enfin nous décrivons le processus d'analyse de l'écriture manuscrite pour détecter la personnalité d'une personne en se basant sur des documents manuscrits.

2.7.1. Travaux connexes

L'identification des traits de personnalité d'un individu à partir de ses archives écrites a fait l'objet de plusieurs travaux de recherche menés par des psychologues contrairement à leurs homologues informatiques qui n'ont pas donné à ce problème d'importance qu'il mérite. Au cours de la dernière décennie, les informaticiens se sont intéressés à ce domaine de recherche, proposant des systèmes automatiques pour automatiser cette tâche.

Champa et coll [39] a présenté une solution qui prédit automatiquement le comportement humain par analyse de l'écriture manuscrite. Ils ont utilisé cinq paramètres d'écriture manuscrite, pression du stylo, ligne de base, hauteur de la lettre « t », boucle inférieure de « y » et inclinaison. Pour l'extraction des

caractéristiques, ils ont utilisé différentes techniques, par exemple, la méthode de polygonation a mise en œuvre pour détecter la ligne de base. Une valeur de seuil de niveau de gris est utilisée pour évaluer la pression du stylo, la correspondance de motif est utilisée pour calculer la hauteur de la lettre « t » et pour l'inclinaison, pour la boucle inférieure de « y » la transformation de Hough généralisée. Après avoir extrait toutes les caractéristiques, ils ont utilisé un système simple basé sur des règles pour prédire les traits de personnalité d'une personne. Dans le cas manuel, la précision moyenne était de 85,8 (%) tandis que la précision de 87,2 (%) était obtenue dans un cas automatisé.

En 2004, MOGHARREBAN, Rahimi et SABHARWAL ont proposé une méthode de prédiction des traits de personnalité basée sur la ligne de base et l'angle d'inclinaison. Sur la base de deux paramètres, les auteurs ont extrait des caractéristiques d'échantillons écrits et ont utilisé le paradigme non diffus (net) et diffus pour évaluer les paramètres de ligne de base et d'angle d'inclinaison. Après avoir obtenu les caractéristiques, ils les ont passées au système Rule-Base pour prédire les différents traits de personnalité, les auteurs ont identifié 15 traits de personnalité [40] .

Une nouvelle méthode d'estimation de l'inclinaison du mot dans le texte écrit à la main a été proposée par RUSTAM LATYPOV et. coll. L'objectif principal de la recherche est de fournir un outil permettant d'observer de petits changements de hauteur d'inclinaison dans les textes pendant le travail. L'extraction de caractéristiques est l'étape la plus importante de l'analyse de l'écriture manuscrite, à l'autre extrémité, une liste des caractéristiques à extraire d'un manuscrit a ensuite été élaborée [41] .

L'année suivante (2015), les auteurs [41] ont utilisé une approche différente de l'apprentissage automatique pour découvrir des traits de personnalité grâce à l'écriture manuscrite. Explorez les différentes caractéristiques de l'écriture manuscrite, y compris la barre en T, la ligne de base, la marge et la pente. Après avoir extrait ces caractéristiques des images numérisées, créé des cartes

de caractéristiques, puis, avec l'aide d'experts sur le terrain (linguistes), ils forment votre modèle. Un algorithme d'apprentissage automatique supervisé (classificateur K-NN) utilisé pour identifier des caractéristiques individuelles à partir des caractéristiques manuscrites extraites.

Toujours dans le même contexte, Joshi P et al ont proposé un outil d'apprentissage automatique qui utilise un classificateur KNN avec apprentissage incrémentiel pour améliorer l'efficacité de l'analyse de l'écriture manuscrite. Les caractéristiques utilisées pour l'analyse de personnalité incluent la détection de la position de base, l'inclinaison des lettres, la hauteur de la barre des lettres et la marge à l'aide de la méthode de polygonation, la marge a été calculée à l'aide de la méthode de balayage vertical, la hauteur de la barre de lettres et la pente des mots ont été calculées à l'aide de la méthode de polygonation.

La plupart de ces caractéristiques sont qualitatives : forme de base, taille des marges, hauteur des lettres, inclinaison du mot.

Le problème de l'observation des caractéristiques humaines est très crucial maintenant, mais il semble assez difficile de révéler tout type de changement de comportement sans un contact direct avec cette personne. L'inclinaison des mots est une caractéristique très importante d'un texte manuscrit. La plupart du temps, les auteurs ont affaire à une détermination qualitative de l'inclinaison : gauche contrôlée, extrême droite, extrême gauche, droite contrôlée et verticale. La précision de la mesure dépend des paramètres de la procédure. Une évaluation est trouvée pour l'inclinaison moyenne des mots, qui s'est avérée être proche d'un angle de 75 degrés. Bien que, cet article reconnaisse la hauteur oblique d'une seule lettre, mais il n'a aucune disposition pour la reconnaissance de l'inclinaison et de la hauteur du mot entier qui, une fois considéré, peut améliorer la reconnaissance d'une image d'entrée [41].

GAVRILESCU le in. Proposé dans [42] une architecture à trois couches pour identifier les traits de personnalité par l'analyse de l'écriture manuscrite. Les auteurs ont analysé les caractéristiques suivantes de l'écriture manuscrite, ligne

de base, inclinaison, pression d'écriture, traits de connexion, lettre minuscule "t", lettre minuscule "f" et espaces entre les lignes. Au niveau de la couche supérieure, un prétraitement est effectué sur l'image numérisée, puis les caractéristiques sont extraites à l'aide de techniques de traitement d'image (par exemple, réduction du bruit, segmentation de phrases, segmentation de mots, etc.) à partir d'un texte manuscrit. Dans la carte d'entités de la couche intermédiaire construite à l'aide de code binaire. Par exemple, le code binaire 010 qui était pour les courses de connexion moyennes.

Dans la dernière couche, un réseau de neurones a été utilisé pour classer les traits. La méthode proposée a atteint une précision de 84,4% pour le test intra-individuel, tandis que pour le test inter-sujet, une précision de 80,5% a été obtenue.

Les chercheurs de [43] ont utilisé l'écriture manuscrite et les caractéristiques de signature pour prédire les traits de personnalité. L'auteur considérait l'espace entre les mots et la ligne, la zone et la ligne de base comme des caractéristiques de l'écriture, tandis que le début de la courbe, la ligne finale, la coque, les lignes intermédiaires, le soulignement, la structure de la période, Séries séparées et déconnectées considérées comme des caractéristiques de l'entreprise. Pour aborder le modèle d'entités simples, les auteurs ont utilisé l'algorithme de structure multiple et, pour les modèles complexes, ils ont utilisé ANN avec MLP. Ils ont extrait des caractéristiques à l'aide de techniques de reconnaissance de formes et dans la dernière étape, ils ont utilisé ANN pour classer les traits de personnalité d'un utilisateur. Grâce aux fonctions d'écriture manuscrite, son système a atteint une précision de 52 à 100% et les fonctions de signature ont atteint une précision de 87 à 100%.

D'autres recherches menées par divers scientifiques dans[44], ont tenté de prédire différents traits de personnalité, par exemple optimiste, pessimiste, équilibré, indépendant et pragmatique, etc.

Sur les échantillons écrits scannés, les auteurs effectuent les étapes de prétraitement. Pour moins de complexité inhérente, les auteurs ont converti l'image en noir et blanc. Dans la dernière étape, CNN est utilisé pour extraire

des caractéristiques de l'échantillon écrit et prédire les différents traits d'un utilisateur.

D'autres recherches menées par différents scientifiques dans [44], ont tenté de prédire différents traits de personnalité, par exemple optimiste, pessimiste, équilibré, indépendant et pragmatique, etc.

Sur les échantillons écrits scannés, les auteurs effectuent les étapes de prétraitement. Pour moins de complexité inhérente, les auteurs ont converti l'image en noir et blanc. Dans la dernière étape, CNN est utilisé pour extraire des caractéristiques de l'échantillon écrit et prédire les différentes caractéristiques d'un utilisateur.

Une enquête plus intéressante par SHITALA et VIVEK dans [45], qui utilise les caractéristiques de l'écriture manuscrite pour détecter les traits de personnalité. Ils ont extrait six types différents d'entités, par exemple, la taille des lettres, l'inclinaison des mots, la ligne de base, la pression du stylet, l'espacement des lettres et les mots écrits. Pour calculer divers paramètres à partir d'échantillons écrits, des techniques de trigonométrie et de seuil sont utilisées. SVM est utilisé dans la dernière étape après le prétraitement et l'extraction des fonctionnalités. Ils ont utilisé le noyau RBF en raison des caractéristiques non linéaires et il a rapporté de meilleurs résultats que le noyau polynomial. Son système a atteint une précision de 93,86%.

Chen et Tao présentent une autre approche pour identifier les traits de personnalité dans [46]. Au total, 23 caractéristiques d'écriture manuscrite ont été prises en compte, par exemple, la taille maximale, minimale et moyenne des lettres, la proportion, les espaces entre les lettres, etc. Pour sélectionner les caractéristiques pertinentes, l'arbre de décision a été utilisé. Ces caractéristiques pertinentes sont alimentées par quatre méthodes de classification différentes les plus couramment utilisées KNN, SVM, ADABOOST et ANN pour analyser les traits de personnalité. Pour ajuster les paramètres optimaux, ils ont utilisé la méthode de recherche GRID sur chaque classificateur. Les auteurs ont noté que les caractéristiques de frappe en ligne

telles que la vitesse de frappe, la pause de frappe et l'accélération contribuent davantage au développement d'une forte personnalité que les caractéristiques hors ligne, en particulier la taille du texte saisi. De toutes les caractéristiques, le décompte rationnel des pauses était la caractéristique d'écriture la plus liée à la personnalité d'une personne. Pour évaluer l'efficacité de la classe, ils ont utilisé des mesures de précision, de sensibilité et de spécificité. Grâce à ces classificateurs, ils ont atteint une précision de (62,5 - 83,9) pour cent.

Dans [47], les auteurs ont introduit une approche de la reconnaissance de la personnalité. En utilisant des caractéristiques indépendantes d'auto corrélation locale d'ordre supérieur (HLAC) et des caractéristiques de contexte, les auteurs ont créé des vecteurs de caractéristiques telles que l'espace entre les tuiles de mots, la taille d'un caractère, l'espace entre la ligne et la marge supérieure. Ils ont utilisé le modèle d'inventaire multiphasique de la personnalité du Minnesota (MMPI) pour l'évaluation.

Le modèle de Markov caché (HMM) est utilisé pour classer les propriétés associées de l'écrivain. Pour les propriétés qui n'étaient pas liées à l'écrivain, MLP est utilisé, leur solution a atteint une précision de 70%.

UMAIR SHAHID a présenté une technique utile pour la prédiction des traits de personnalité grâce à l'analyse informatisée de l'écriture manuscrite. La méthode repose sur les caractéristiques de l'écriture manuscrite. L'extraction des caractéristiques et la classification des traits de personnalité sont effectuées à l'aide du réseau neuronal convolutif. Le modèle à cinq facteurs est utilisé contre l'évaluation des résultats. Dans la première étape, les images numérisées sont converties en niveaux de gris. Dans la deuxième étape, des fonctionnalités extraites d'images numérisées à l'aide de CNN et à l'étape finale, un réseau de neurones est formé par rétro propagation pour apprendre les modèles à partir des fonctionnalités extraites et déterminer les traits de personnalité des Big Five. Pour former et tester l'architecture proposée, nous créons d'abord une telle base de données (en raison de l'indisponibilité de l'ensemble de données manuscrites, qui relie les traits FFM à des échantillons d'écriture manuscrite) qui se compose de 285 échantillons écrits.

Le résultat du test sur des échantillons écrits au hasard a obtenu une précision de prédiction globale de 59,73% [48].

2.8. Analyse critique

Il a été observé que différents chercheurs utilisaient différentes techniques dans cet état de l'art. La ligne de base et l'inclinaison des mots étaient les caractéristiques les plus couramment utilisées par presque tous les scientifiques pour prédire les traits de personnalité - un minimum de 56 et un maximum de 128 échantillons écrits utilisés dans ce type d'étude. Un point supplémentaire considéré ici en ce qui concerne l'extraction de caractéristiques, sauf LEMOS et coll [44] chaque auteur a utilisé des caractéristiques faites à la main pour identifier les traits. Ils ont besoin d'une certaine connaissance du domaine ou d'un expert pour traiter les caractéristiques du texte écrit, et c'était le problème.

Bien que certains auteurs aient utilisé des techniques d'apprentissage automatique pour obtenir d'excellents résultats. Même ainsi, les auteurs n'ont pas validé l'exactitude du modèle, à l'exception de GAVRILESCU et coll [42] et Chen et coll [46] leurs résultats sont authentiques car ils ont confirmé les traits de personnalité prédits par leur modèle avec les résultats du modèle à cinq facteurs, mais ils ont également utilisé des caractéristiques faites à la main.

9. Conclusion

Dans ce chapitre, nous avons présenté dans un premier temps, quelques domaines de recherche qui sont en liaison directe avec le problème étudié, ces domaines incluent : l'authentification de signatures manuscrite, la reconnaissance de l'écriture manuscrite et l'identification de scripteurs à partir de documents manuscrits, nous nous sommes concentré ensuite sur les travaux connexes dans le domaine de l'identification des traits de personnalité à partir d'images de textes manuscrits, puis, nous avons décrit les cinq grands traits de personnalité considérés dans le cadre de ce mémoire. Après avoir présenté une étude détaillée sur caractéristiques de l'écriture manuscrite et la personnalité. Nous avons terminé ce chapitre par une conclusion précédée par la

présentation du processus d'analyse de l'écriture manuscrite pour détecter la personnalité d'une personne en se basant sur des documents manuscrits.

Caractéristiques texturales pour l'identification de traits de personnalité à partir de documents manuscrits

Ce chapitre présente notre contribution principale qui consiste en une étude sur l'identification de traits de personnalité à partir de documents manuscrits. La méthode proposée est basée sur l'extraction d'un ensemble de caractéristiques texturales à partir d'échantillons de documents manuscrits et l'entraînement d'un classifieur afin qu'il puisse identifier les traits de personnalité du scripteur d'un document en question. Des attributs comme l'orientation, la courbure et la texture sont estimés en calculant les distributions des charnières des contours (EDGE-HINGE). La classification est effectuée à l'aide des forêts aléatoires (RF), les Machines à vecteurs de support (SVM) ainsi qu'un algorithme de booster adaptatif (ADABOOST). La méthode proposée a été évaluée en utilisant une base de données contenant 285 documents manuscrits où des résultats intéressants ont été enregistrés.

3.1. Introduction

Les traits de personnalité sont d'une importance capitale pour identifier le comportement de l'être humain. Ils représentent une source d'information utile pour le contrôle médico-légal, le profilage du recrutement, les symptômes médicaux et d'autres applications. Les traits de personnalité sont identifiés à travers divers aspects physiques, y compris le sens, l'honnêteté et d'autres émotions. Ces aspects peuvent être révélés grâce à des attributs de l'écriture manuscrite. Puisque l'écriture est unique pour tout le monde, son processus d'identification n'est pas aussi simple tel qu'il apparaît, il s'agit plutôt d'outils efficaces pour l'extraction et la classification des caractéristiques.

Le processus a fait l'objet de divers travaux de recherche. Cependant, les résultats rapportés restent insatisfaisants en raison principalement de la variabilité des écritures manuscrites.

Dans ce chapitre, nous présentons une méthode d'identification des traits de personnalité basée sur l'estimation d'attributs comme l'orientation, la courbure et la texture en calculant les distributions des charnières des contours (EDGE-HINGE). La classification est effectuée à l'aide des forêts aléatoires (RF), les Machines à vecteurs de support (SVM) ainsi qu'un algorithme de booster adaptatif (ADABOOST). La méthode proposée a été évaluée en utilisant une base de données contenant 285 documents manuscrits où des résultats intéressants ont été enregistrés.

Ce chapitre est organisé comme suit. Dans la section suivante, nous décrivons la base de données utilisée. Nous présentons ensuite la description de la méthode de détection et d'élimination des lignes de portée choisie ainsi que la technique de classification utilisée dans les sections 3.3 et 3.4 respectivement, suivies par la description des caractéristiques proposées dans la section 3.5. Les résultats expérimentaux et leur analyse sont présentés dans la section 3.6, tandis que la dernière section conclut ce chapitre.

3.2. Ensemble de données

L'ensemble de données utilisé pour l'évaluation du présent travail [48] consiste en un nouveau corpus, dédié à l'expérience du problème des traits de personnalité. Son contenu se compose de questionnaires FFM et d'échantillons d'écriture manuscrite d'un groupe de 285 sujets (hommes et femmes), âgés de 20 à 40 ans. Pour chaque sujet, deux paragraphes sont disponibles: Le texte du premier paragraphe était le même pour tous les participants tandis que le second était différent pour chaque personne. Certains échantillons manuscrits sont illustrés à la figure 3.1.

This is a humanitarian crisis. We will do everything possible to help these people. If the president will not provide us with financial assistance to house and feed these people, he will be creating a homeless encampment," Mayor Bogen said in a statement, adding that "a sudden influx of immigrants will further strain Broward County's services and will cause further harm to immigrants who will be left here with no money, housing or basic knowledge of the area." Bogen said the incoming migrants will be processed at U.S. Customs and Border Protection Offices in the two counties in Davia Beach and West Palm Beach.

This is a humanitarian crisis. We will do everything possible to help these people. If the president will not provide us with financial assistance to house and feed these people, he will be creating a homeless encampment," Mayor Bogen said in a statement, adding that "a sudden influx of immigrants will further strain Broward County's social services and will cause further harm to immigrants who will be left here with no money, housing or basic knowledge of the area. Bogen said the incoming migrants will be processed at U.S. Customs and Border Protection Office in the two counties in Davia Beach and West Palm Beach.

Figure 3.1: Échantillons manuscrits extraits de l'ensemble de données utilisé.

[48]

Un identifiant unique est attribué à chaque formulaire d'enquête pour garder une trace des résultats FFM et des échantillons manuscrits. Différents résultats sont obtenus pour chaque utilisateur après avoir rempli le questionnaire FFM, et la plage de résultats pour chaque trait était comprise entre 0 et 100. Les résultats de chaque trait ont été classés en trois types tels que faible (L), moyen (M) et élevé (H). Par exemple, si la valeur du trait est inférieure à 33, une étiquette (L) lui est attribuée ou si une valeur supérieure ou égale à 34

mais inférieure ou égale à 66 alors l'étiquette (M) lui est affectée et si la valeur est supérieure ou égale à 67, l'étiquette (H) lui est assignée . Le regroupement des résultats des traits est expliqué à la figure 3.2.

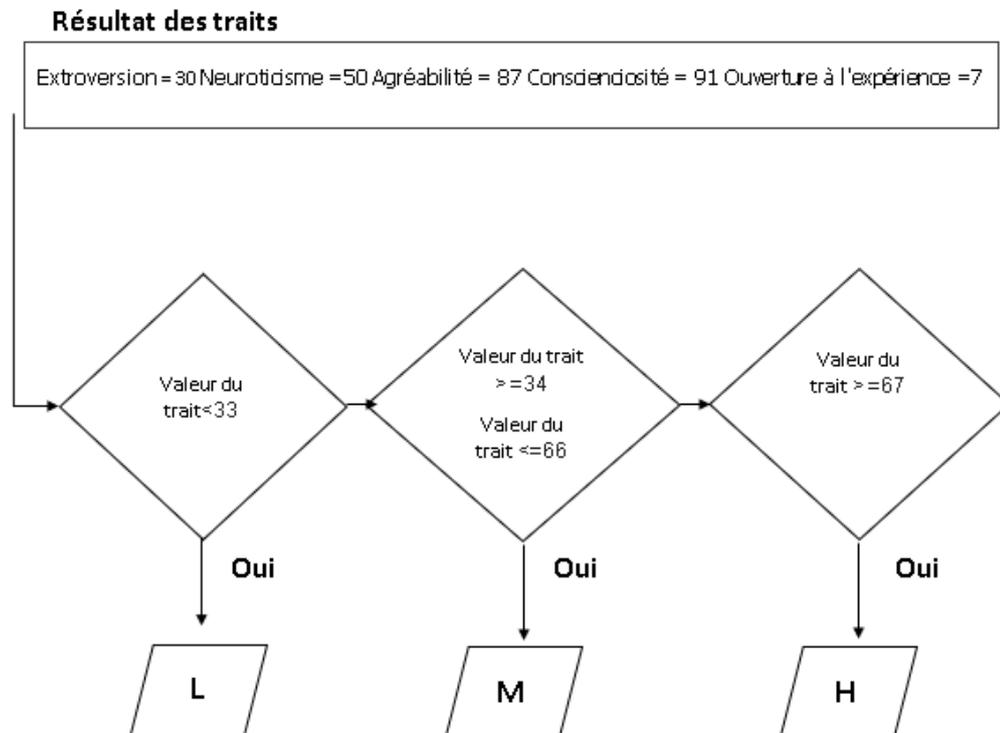


Figure 3.2: Catégorisation des résultats FFM. [48]

Il est à noter que chaque paragraphe est étiqueté avec cinq classes correspondant aux cinq grands traits de personnalité. Ces traits sont : l'extraversion, le névrosisme, la conscienciosité , l'agrément, et l'ouverture à l'expérience.

Le nombre total d'échantillons dans le corpus utilisé est de 570, il est divisé en trois sous-ensembles: apprentissage, validation et test. L'ensemble de validation est utilisé pour le réglage de paramètres du système si nécessaire, tandis que le sous ensemble de test est utilisé pour l'évaluation des performances du système proposé. Le tableau 3.1 montre la répartition des scripteurs (également appelés sujets) par trait de personnalité.

Trait	Nombre d'échantillons		
	Faible (L)	Moyen (M)	Elevé (H).
Extraversion	152	310	108
Névrosisme	174	306	90
Conscienciosité	214	274	82
Agréabilité	218	278	74
Ouverture à l'expérience	402	134	34
Total	1160	1302	389

Tableau 3.1: Répartition des échantillons sur les cinq traits de personnalité.

3.3. Extraction de caractéristiques

Le système d'identification de traits de personnalité proposé est basé sur un ensemble d'attributs comme l'orientation, la courbure et la texture sont estimés en calculant les distributions des charnières des contours (Edge-HINGE) [49].

Afin de capturer la courbure du contour, ainsi que son orientation, les distributions des charnières des contours sont utilisées. L'idée centrale est de considérer dans le voisinage, non pas un, mais deux fragments de bord émergeant du pixel central et de calculer ensuite la distribution de probabilité conjointe des orientations des deux fragments. Nous avons considéré dans toutes les expérimentations qui seront présentées dans les sections suivantes des fragments de tailles allant de 4 pixels à 9 pixels. La figure 3.3 présente une description graphique des caractéristiques des charnières des contours.

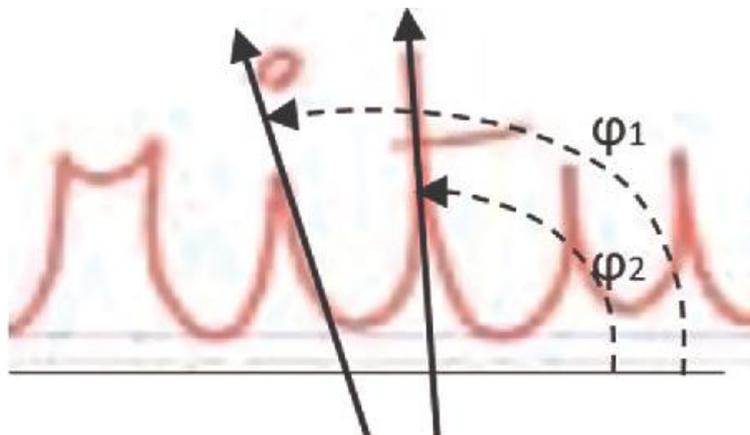


Figure 3.3 : Description graphique des caractéristiques des charnières des contours.
[55]

Le tableau 3.2 résume, pour chacune des caractéristiques utilisées, le numéro correspondant, la description et la dimension.

Caractéristique	Description	Dimension
<i>f1</i>	Distribution des charnières des contours avec un fragment de 4 pixels	576
<i>f2</i>	Distribution des charnières des contours avec un fragment de 5 pixels	1024
<i>f3</i>	Distribution des charnières des contours avec un fragment de 6 pixels	1600
<i>f4</i>	Distribution des charnières des contours avec un fragment de 7 pixels	2304
<i>f5</i>	Distribution des charnières des contours avec un fragment de 8 pixels	3136
<i>f6</i>	Distribution des charnières des contours avec un fragment de 9 pixels	4096

Tableau 3.2 : Une vue d'ensemble des caractéristiques implémentées et leurs dimensionnalités.

3.4. Classification

Une fois que les images de documents à comparer sont représentées par leurs caractéristiques, nous procédons à l'utilisation de ces vecteurs de caractéristiques pour la classification. L'entraînement et la classification sont effectués à l'aide des forêts aléatoires (RF), les Machines à vecteurs de support (SVM) ainsi qu'un algorithme de boosting adaptatif (ADABOOST) en utilisant les ensembles de données d'apprentissage et de test tandis que les différents paramètres des différents classifieurs sont déterminés empiriquement sur l'ensemble des données de validation. Une brève introduction à ces classifieurs est donnée dans les paragraphes suivants.

3.4.1. Les forêts aléatoires (RF)

Un classificateur de forêts aléatoires [50] se compose d'un certain nombre d'arbres, chaque arbre étant cultivé en utilisant une forme de randomisation. Les nœuds feuilles de chaque arbre sont étiquetés par des estimations de la distribution postérieure sur les classes d'images de documents manuscrits. Chaque nœud interne contient un test qui divise au mieux l'espace de données à classer.

Une image d'un document manuscrit est classée en l'envoyant dans chaque arbre et en agrégeant les distributions de feuilles atteintes. L'aléatoire peut être injecté en deux points pendant l'apprentissage; en sous-échantillonnant les données d'apprentissage afin que chaque arbre soit développé en utilisant un sous-ensemble différent; et dans la sélection des tests de nœuds.

3.4.2. Les Machines à vecteurs de support (SVM)

La méthode SVM est une méthode de classification linéaire [51] qui repose sur l'hypothèse que, étant donné un espace approprié, il existe un classificateur linéaire (appelé hyperplan) permettant de distinguer les deux classes de l'espace (+/-). Le but de cette méthode est d'apprendre, à partir d'un ensemble d'exemples d'apprentissage (apprentissage supervisé), une fonction qui prédit les classes pour de nouveaux objets. Plus concrètement, il s'agit de trouver l'hyperplan optimal, qui sépare les données et maximise la distance entre les deux classes.

L'hyperplan optimal est celui, parmi tous les hyperplans valides, qui réalise la marge maximale entre les points des deux classes. C'est la raison pour laquelle on parle de séparateur à vaste marge. Les points les plus proches de la frontière entre les deux classes et qui sont utilisés pour la détermination de l'hyperplan optimal sont appelés vecteurs supports. L'hyperplan optimal est celui qui permettra au mieux de classer les nouveaux exemples. La classification d'un nouvel exemple inconnu est donnée par sa position par rapport à l'hyperplan optimal. Face à un cas non linéairement séparable (c'est-à-dire la plupart des problèmes réels), les méthodes SVM recourent à une fonction

noyau pour effectuer une transformation non linéaire des données. Le résultat de cette transformation, appelé espace de re-description, est un espace de dimension plus grande.

3.4.3. Algorithme de boosting adaptatif (ADABOOST)

ADABOOST [52] est un algorithme d'apprentissage ensembliste avec qui consiste à résoudre des problèmes de classifications complexes en intégrant des classifieurs faibles simples à des classifieurs forts. Les modèles d'apprentissage ensemblistes sont conçus selon deux approches, BOOSTING et Bagging. ADABOOST est l'un des algorithmes de BOOSTING où les poids de l'instance individuelle sont déterminés de manière itérative en se basant sur la précision du dernier résultat de classification. C'est un classifieur très précis qui offre un taux d'erreur proche de zéro.

ADABOOST semble être sensible aux données bruitées et les valeurs aberrantes. ADABOOST doit subir deux actions, premièrement, le classifieur est entraîné de manière itérative sur diverses données d'apprentissage pondérées. Deuxièmement, à chaque itération, il fournit un bon ajustement pour les instances en minimisant l'erreur d'apprentissage. De nombreux classifieurs sont combinés par L'ADABOOST pour augmenter leurs précisions.

3.5. Résultats et discussion

Nous présentons et analysons les performances des caractéristiques proposées dans l'identification des traits de personnalité. La mesure de performances utilisée est le taux d'identification. Toutes les expérimentations sont menées sur l'ensemble de données présenté dans la section 3.2. Nous avons envisagé quatre scénarios d'évaluation différents en considérant les trois classifieurs présentés dans la section 3.4.

Nous avons effectué plusieurs séries d'expérimentation dans l'objectif de déterminer les meilleurs paramètres du classifieur utilisé conduisant aux meilleures performances. Lors de toutes les séries d'expérimentation, La

première image de chacun des 285 scripteurs est utilisée pour l'apprentissage tandis que l'autre est utilisée dans le test.

Nous avons envisagé quatre scénarios d'évaluation différents. Dans le premier, nous avons calculé les performances globales des différentes caractéristiques en utilisant les trois classifieurs (RF, SVM et ADABOOST). Dans les trois scénarii suivants, nous avons calculé les performances des différentes caractéristiques pour les différents traits de personnalités en utilisant les trois classifieurs (RF, SVM et ADABOOST).

Le tableau 3.3 présente les taux globaux d'identification enregistrés en utilisant les forêts aléatoires (RF), les Machines à vecteurs de support (SVM) ainsi que l'algorithme de boosting adaptatif (ADABOOST) en se basant sur les caractéristiques considérées.

<i>Caractéristique</i>	<i>Taux d'identification global</i>		
	RF	SVM	AdaBoost
<i>f1</i>	<i>50.17%</i>	<i>48.88%</i>	<i>82.25%</i>
<i>f2</i>	<i>49.94%</i>	<i>54.50%</i>	<i>82.29%</i>
<i>f3</i>	<i>50.76%</i>	<i>54.50%</i>	<i>79.29%</i>
<i>f4</i>	<i>49.59%</i>	<i>54.50%</i>	<i>82.10%</i>
<i>f5</i>	<i>49.00%</i>	<i>54.50%</i>	<i>77.66%</i>
<i>f6</i>	<i>49.82%</i>	<i>54.50%</i>	<i>81.40%</i>

Tableau 3.3 : Taux globaux d'identification enregistrés en utilisant les forêts aléatoires (RF), les Machines à vecteurs de support (SVM) ainsi que l'algorithme de BOOSTING adaptatif (ADABOOST).

Les tableaux 3.4, 3.5 et 3.6 présentent les résultats réalisés pour les cinq traits de personnalités en utilisant les caractéristiques proposées pour les classifieurs considérés.

<i>Caractéristique</i>	Extraversion	Neuroticisme	Agréabilité	Conscienciosité	Ouverture à l'expérience
<i>f1</i>	44.44%	43.85%	42.10%	69.59%	68.42%
<i>f2</i>	49.70%	43.85%	42.69%	45.02%	68.42%
<i>f3</i>	49.12%	43.85%	43.85%	46.78%	70.17%
<i>f4</i>	45.61%	43.85%	40.93%	47.36%	70.17%
<i>f5</i>	43.27%	43.85%	42.69%	45.61%	69.59%
<i>f6</i>	46.78%	43.85%	42.69%	47.95%	67.83%

Tableau 3.4 : Résultats réalisés pour les cinq traits de personnalités en utilisant les forêts aléatoires (RF).

<i>Caractéristique</i>	Extraversion	Neuroticisme	Agréabilité	Conscienciosité	Ouverture à l'expérience
<i>f1</i>	26.90%	43.85%	43.85%	55.55%	74.26%
<i>f2</i>	54.97%	43.85%	43.85%	55.55%	74.26%
<i>f3</i>	54.97%	43.85%	43.85%	55.55%	74.26%
<i>f4</i>	54.97%	43.85%	43.85%	55.55%	74.26%
<i>f5</i>	54.97%	43.85%	43.85%	55.55%	74.26%
<i>f6</i>	54.97%	43.85%	43.85%	55.55%	74.26%

Tableau 3.5 : Résultats réalisés pour les cinq traits de personnalités en utilisant les machines à vecteurs de support (SVM).

<i>Caractéristique</i>	Extraversion	Neuroticisme	Agréabilité	Conscienciosité	Ouverture à l'expérience
<i>f1</i>	90.64%	83.04%	83.04%	82.45%	73.68%
<i>f2</i>	85.96%	84.79%	84.79%	82.45%	76.60%
<i>f3</i>	88.88%	78.94%	78.94%	82.45%	67.25%
<i>f4</i>	89.47%	81.28%	81.28%	82.45%	76.02%
<i>f5</i>	84.79%	78.94%	78.94%	77.77%	67.83%
<i>f6</i>	83.04%	85.38%	85.38%	80.11%	73.09%

Tableau 3.6 : Résultats réalisés pour les cinq traits de personnalités en utilisant l'algorithme de boosting adaptatif (ADABOOST).

Une analyse des résultats présentés dans les tableaux 3.3, 3.4, 3.5 et 3.6 conduit aux observations suivantes :

- ✚ Les distributions des charnières des contours avec un fragment de 5 pixels (f2) réalisent les meilleurs taux globaux d'identification avec l'algorithme de boosting adaptatif (ADABOOST).
- ✚ Les taux globaux d'identification enregistrés en utilisant les forêts aléatoires (RF) et les machines à vecteurs de support (SVM) ne sont pas performants si on les compare à ceux réalisés avec l'algorithme de boosting adaptatif (ADABOOST).
- ✚ L'ouverture à l'expérience est le trait le mieux caractérisé par toutes les caractéristiques considérées en utilisant les forêts aléatoires (RF) et les machines à vecteurs de support (SVM), alors que l'extraversion est le trait le mieux caractérisé par l'algorithme de boosting adaptatif (ADABOOST) en utilisant les mêmes caractéristiques.
- ✚ Enfin, les résultats obtenus démontrent clairement le potentiel des caractéristiques proposées ainsi que l'algorithme de boosting adaptatif (ADABOOST) pour l'identification des traits de personnalité à partir d'images scannées de l'écriture manuscrites.

Nous présentons également une comparaison de notre système avec les systèmes de pointe des quatre dernières années rapportées dans la littérature. Les comparaisons sont résumées dans le tableau 8.

Study	Classifieur	Dataset	Résultats
Fallah et al [10] (2016)	HMM and ANN	70 Subjects	70%
Gavrilescu et al. (2018) [11]	ANN	128 Subjects	80.5 %
UMAIR SHAHID[2] (2020)	CNN	285 Subjects	59,73%
Technique proposée (2021)	ADABOOST	285 Subjects	83%

Tableau 3.7 : Comparaison des performances des systèmes de prédiction des traits de personnalité basés sur l'écriture manuscrite .

2.5. Conclusion

Ce travail avait pour objectif de présenter une méthode pour l'identification de traits de personnalité à partir de documents manuscrits. Nous avons utilisé un ensemble de caractéristiques texturales qui ont montré des résultats très prometteurs sur une base de données de documents manuscrits. Les évaluations ont été effectuées sur une base de données contenant des échantillons de textes manuscrits collectés auprès de 285 sujets différents. Les résultats obtenus reflètent l'efficacité des caractéristiques proposées. Un autre aspect intéressant de cette étude est l'évaluation d'un certain nombre de classifieurs sur cette base de données. Dans tous les cas, certains classifieurs se sont montrés plus performants que d'autres classifieurs.

La contribution que nous avons proposée dans le cadre de ce mémoire nous ont permis d'aboutir à des résultats prometteurs, mais nous ont aussi ouvert plusieurs voies pouvant être exploitées dans le futur. Les études ultérieures sur ce sujet seront destinées à introduire des caractéristiques supplémentaires et ensuite appliquer un mécanisme de sélection de caractéristiques pour savoir quelles sont les caractéristiques les plus discriminantes pour ce problème et pour des problèmes similaires. Il est nécessaire de rappeler que la performance du système proposé ne dépend pas seulement des techniques de classification utilisées, mais aussi des caractéristiques choisies.

Dans ce cadre, il serait très intéressant d'exploiter la combinaison de caractéristiques proposées dans ce mémoire avec celles de l'état de l'art afin d'améliorer les performances du système proposé. Pour les techniques de classification utilisées, nous pensons qu'il serait intéressant d'envisager l'utilisation d'autres techniques de classification que celles que nous avons adoptées dans le présent mémoire. Il serait aussi très intéressant aussi de considérer des bases de données plus volumineuses que celle utilisée dans le cadre du présent travail.

CONCLUSION

GENERALE

Ce travail a abordé le problème de l'identification des traits de personnalités à partir de documents manuscrits en utilisant les distributions des charnières des contours (Edge-HINGE) comme caractéristiques. Les forêts aléatoires (RF), les machines à vecteurs de support (SVM) ainsi qu'un algorithme de boosting adaptatif (ADABOOST) ont été employés pour la classification.

Des échantillons de documents manuscrits issus de 258 scripteurs différents ont été utilisés pour l'évaluation des performances du système proposé. Le système proposé a réalisé des performances très encourageantes en particulier lors de l'utilisation de l'algorithme de boosting adaptatif (ADABOOST).

Les études ultérieures sur ce sujet seront destinées à introduire des caractéristiques supplémentaires et ensuite appliquer un mécanisme de sélection de caractéristiques pour savoir quelles sont les caractéristiques les plus discriminantes pour ce problème et pour des problèmes similaires. Il est nécessaire de rappeler que la performance du système proposé ne dépend pas seulement des techniques de classification utilisées, mais aussi des caractéristiques choisies.

Dans ce cadre, il serait très intéressant d'exploiter la combinaison de caractéristiques proposées dans ce mémoire avec celles de l'état de l'art afin d'améliorer les performances du système proposé. Pour les techniques de classification utilisées, nous pensons qu'il serait intéressant d'envisager l'utilisation d'autres techniques de classification que celles que nous avons adoptées dans le présent mémoire. Il serait aussi très intéressant aussi de considérer des bases de données plus volumineuses que celle utilisée dans le cadre du présent travail.

Bibliography

- [1] Janet Fisher, Anish MAREDIA, Anita Nixon, Nerissa Williams, and Jonathan LEET. Identifying personality traits, and especially traits resulting in violent behavior through automatic handwriting analysis. Proceedings of Student-Faculty Research Day, CSIS, Pace University, 2012.
- [2] Nikita LEMOS, KRISH Shah, Rajas RADE, AND DHARMIL Shah. Personality prediction based on handwriting using machine learning. In 2018 International Conference on Computational Techniques, Electronics and Mechanical Systems (CTEMS), pages 110–113. IEEE, 2018.
- [3] SALANKARA Mukherjee and Ishita De. Feature extraction from handwritten documents for personality analysis. In 2016 International Conference on Computer, Electrical & Communication Engineering (ICCECE), pages 1–8. IEEE, 2016.
- [4] Manoj SACHAN HEMLATA and Shailendra Kumar Singh. Personality detection using handwriting analysis.
- [5] Anamika Sen and Harsh Shah. Automated handwriting analysis system using principles of graphology and image processing. In 2017 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS), pages 1–6. IEEE, 2017.
- [6] KUKUH ADI PRASETYO, NANA RAMADIJANTI, and ACHMAD Basuki. Mobile application for identifying personality of person using graphology. In 2017 International Electronics Symposium on Knowledge Creation and Intelligent Computing (IES-KCIC), pages 212–219. IEEE, 2017.
- [7] ZHI Chen and Tao Lin. Automatic personality identification using writing BEHAVIOURS: an exploratory study. BEHAVIOUR & Information Technology, 36(8):839–845, 2017.
- [8] Wei XIONG. An investigation of big five and narrow personality traits in relation to career satisfaction of managers. 2010.
- [9] Seema KEDAR, Vaishnavi Nair, and Shweta Kulkarni. Personality identification through handwriting analysis: a review. Int. J. Adv. Res. COMPUT. Sci. SOFTW. ENG, 5(1), 2015.
- [10] B Baggett. Handwriting analysis 101-the basic traits. Published by EMPRESSE Publishing Distributed by Handwriting University, 2004.
- [11] Afran H GAROOT, MAEDEH Safar, and Ching Y Suen. A comprehensive survey on handwriting and computerized graphology. In 2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR), volume 1, pages 621–626. IEEE, 2017.
- [12] Ricard Coll, Alicia FORNÉS, and JOSEP LLADÓS. Graphological analysis of handwritten text documents for human resources recruitment. In 2009 10th International Conference on Document Analysis and Recognition, pages 1081–1085. IEEE, 2009.
- [13] Esmeralda C DJAMAL, SHELDY NUR RAMDLAN, JERI SAPUTRA Recognition of Handwriting Based on Signature and Digit. Information Systems International Conference (ISICO), December 2013.
- [14] K Amend et al. The Complete Basic Book 1981 Handwriting Analysis: The Complete Basic Book, California.
- [15] G. LOULLOUDIS et al, Line and Word Segmentation of Handwritten.

- [16] Panwar, N. Nain and S. Handwritten text recognition system based on neural network.
- [17] Markus Jokela, Christian HAKULINEN, Archana Singh-MANOUX, and Mika KIVIMÄKI
Personality change associated with chronic diseases *Psychological Medicine*, 44 ,2014.
- [18] MROCZEK, Brent W Roberts and Daniel, Personality trait change in adulthood.2008,
Current directions in psychological science.
- [19] VIJAYALATA, KASULA CHAITHANYA Pramodh and al. Automatic personality recognition of authors using big five factor model, *International Conference on Advances in Computer Applications (ICACA) 2016*.
- [20] John, Robert R McCrae and Oliver P, An introduction to the five-factor model and its applications,1992. *Journal of personality*.
- [21] Soto., Joshua J Jackson and Christopher J, Five-factor model of personality. *Oxford bibliographies*.
- [22] John et Srivastava 1999.
- [23] LEBOWIT 2016.
- [24] Timothy Judge, Jessica RODELL, Ryan Klinger, Lauren Simon, and E.EAN Crawford, Hierarchical Representations of the Five-Factor Model of Personality in Predicting Job Performance: Integrating Three Organizing Frameworks with Two Theoretical Perspectives.2013
- [25] ANANDAKUMAR, HN CHAMPA and KR, Automated human behavior prediction through handwriting analysis, *First International Conference on Integrated Intelligent Computing*.2010
- [26] HUWIDA E. Said, K.D. Baker, Writer identification based on handwriting,1998 *Handwriting Analysis and Recognition (Ref. No. 1998/440)*, IEE Third European Workshop on UK.
- [27] N MOGHARREBAN, Shahram Rahimi, and M Sabharwal, A combined crisp and fuzzy approach for handwriting analysis. *IEEE Annual Meeting of the Fuzzy Information*, 2004.
- [28] GATTAL ABDELJALIL, CHAWKI DJEDDI, IMRAN SIDDIQI, SOMAYA AL-MAADEED, Writer Identification on Historical Documents Using Oriented Basic Image Features, *16th International Conference on Frontiers in Handwriting Recognition*,2018.

[29] Ricardo GRAÇA, Rui SARMENTO e Castro, and João CEVADA, PARKDETECT: Early diagnosing, IEEE International Symposium on Medical Measurements and Applications (MEMEA) 2014.

[30] OURAGHE Ibtissame¹, AMMOUR Alae¹, ABOULEM Ghita², KHAISSIDI Ghizlane¹, MRABTI Mostafa¹, BELAHSEN Faouzi², MOUNIM A. El-Yacoubi³, Sonia Garcia-Salicetti³, ANALYSE AUTOMATIQUE DE L'ÉCRITURE MANUSCRITE EN LIGNE POUR LA DÉTECTION PRÉCOCE DES PATHOLOGIES NEURODÉGÉNÉRATIVES, Automatic analysis of on-line handwriting for the early detection of neurodegenerative pathologies.2017

[31] Sara Rosenblum, BATYA ENGEL-YEGER, and Yael Fogel, Age-related changes in executive control and their relationships with activity performance in handwriting, Human movement science,2013.

[32] LOSTEC, Claire LE, Les troubles du graphisme et l'écriture chez l'enfant dyspraxique, CONNAISSANCES SURDITÉS 2008.

[33] SV KEDAR, DS BORMANE, AADITI DHADWAL, SHIWALI Alone, and RASHI AGARWAL, Automatic emotion recognition through handwriting analysis, International Conference on Computing Communication Control and Automation,2015.

[34] LAURENCE LIKFORMAN-SULEM, ANNA ESPOSITO, MARCOS FAUNDEZ-ZANUY, STEPHAN CLÉMENÇON, GENNARO CORDASCO, EMOTHAW: A Novel Database for Emotional STATE, IEEE.

[35] Florence Brun-Henin, Jean-Luc Velay, Yaël Beecham, Sophie CARIOUTROUBLES d'écriture et dyslexie : revue théorique, aspects cliniques et approche expérimentale,2012.

[36] LAURENCE LIKFORMAN-SULEM, Anna Esposito, Marcos FAUNDEZ-ZANUY, and STÉPHAN CLÉMENÇON, EXTRACTING style and emotion from HANDWRITING, IN Advances in Neural Networks: Computational and Theoretical Issues,2015.

[37] Thiry, Benjamin, Écriture manuscrite et personnalité : évaluation de la fiabilité et de la validité de la graphologie,2021.

[38] N MOGHARREBAN, Shahram Rahimi, and M SABHARWAL, A combined crisp and fuzzy approach for handwriting analysis, IEEE Annual Meeting of the Fuzzy Information,2004.

[39] ANANDAKUMAR, HN CHAMPA and KR, AUTOMATED human behavior prediction through handwriting analysis, In 2010 First International Conference on Integrated Intelligent Computing,2010.

- [40] N MOGHARREBAN, Shahram Rahimi, and M Sabharwal. A combined crisp and fuzzy approach for handwriting analysis, In IEEE Annual Meeting of the Fuzzy Information, 2004.
- [41] Prachi Joshi, Aayush Agarwal, Ajinkya DHAVAL, Rajani SURYAVANSHI, and Shreya KODOLIKAR, International Journal of Computer Applications,2015.
- [42] Marcus LIWICKI, Andreas SCHLAPBACH, Peter LORETAN, and Horst BUNKE, Automatic detection of gender and handedness from on-line handwriting,13th Conf. of the GRAPHONOMICS Society ,2007.
- [43] Esmeralda Contessa DJAMAL, RISNA DARMAWATI, and SHELDY Nur RAMDLAN, Application image processing to predict personality based on structure of handwriting and signature, International Conference on Computer, Control, Informatics and Its Applications (IC3INA),2013.
- [44] Nikita LEMOS, KRISH Shah, Rajas RADE, and DHARMIL Shah, Personality prediction based on handwriting using machine learning, International Conference on Computational Techniques, Electronics and Mechanical Systems (CTEMS) 2018.
- [45] SHITALA Prasad, Vivek Kumar Singh, and AKSHAY SAPRE, segmentation method for prediction of human personality using support vector machine, 2010International Journal of Computer Applications
- [46] Lin, ZHI Chen and Tao, Automatic personality identification using writing BEHAVIOURS: an exploratory study, BEHAVIOUR & Information Technology,2017.
- [47]H. KHOTANLOU, «Identify human personality parameters based on handwriting using neural network» chez Artificial Intelligence and Robotics (IRANOPEN), 2016.
- [48]SHAHID, UMAIR, IDENTIFICATION OF PERSONALITY TRAITS USING COMPUTERIZED ANALYSIS OF HANDWRITING,2020
- [49](C) Laurens van der MAATEN,Université Maastricht, 2007.
- [50] Pal, M, RANDOM forest classifier for remote sensing classification, International Journal of Remote Sensing 2003.
- [51] SURYANNARAYANA CHANDAKA, AMITAVA CHATTERJEE, Cross-correlation aided support vector machine classifier for classification of EEG signals,ELSEVIER 2009.

[52] SCHAPIRE, Yoav Freund Robert E, Experiments with a New Boosting Algorithm, AT&T Research,1996.

[53] Handwriting Analysis Chart: Handwriting Analysis Traits,

[https://www.handwriting-graphology.com/handwriting-analysis-chart/\(05/04/2021\)](https://www.handwriting-graphology.com/handwriting-analysis-chart/(05/04/2021))

[54]The Big Five Personality Traits Model and Test: Using OCEAN to Match Roles to Characteristics (05/04/2021),https://www.mindtools.com/pages/article/newCDV_22.htm

[55]CHAWKI DJEDDI, TAHAR MEKHAZANIA, SOBHAN SARKAR,Personality Traits Identification Through Handwriting Analysis, March 2021.

[56]S. KEDAR, D. BORMANE, AADITI DHADWAL, SHIWALI ALONE, RASHI AGARWAL, Automatic Emotion Recognition through Handwriting Analysis: A Review, 2015 International Conference on Computing Communication Control and Automation.