

N°d'ordre :
N° de série :

RÉPUBLIQUE ALGÉRIENNE DÉMOCRATIQUE ET POPULAIRE
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique



UNIVERSITE ECHAHID HAMMA LAKHDAR - EL OUED
FACULTÉ DES SCIENCES EXACTES
Département D'Informatique



Mémoire de Fin D'étude
Présenté pour l'obtention du Diplôme de

LICENCE ACADEMIQUE

Domaine : **Mathématique et Informatique**
Filière : **Informatique**
Spécialité : **Systemes Informatiques**

Thème

Conception et développement d'un application de classification des image de document

Présenté par :

- **OTHMANI Oum Elhana**
- **NEDJAI Ahlam**
- **BOUAMRA Richa**

Proposé et Encadré par : **M. KHEBBACHE Mohib Eddine**

Soutenue le xx-xx- 2018 Devant le jury:

M.

Président

M.

Rapporteur

Année Universitaire: 2017-2018

DÉDICACE

Je tiens à dédier ce modeste travail à la
lumière de ma vie : mes chers parents , mon
père Ali , et ma mère Khaira pour leur plus
grand amour ,

Soutien , encouragement de la patience et de
l'aide contenue pendant mes années d'études .

Ce travail est également dédié :

A mes soeurs douce Awatef,Dhouha,Arwa .

Pour ma cher frère Mohammed .

Pour tous mes oncles et tantes

Mes dédicace vont tendrement à mes chère
amies : Hana , chaima , Safia ,Youmna ,

Khadija, Tayeb , Ziad, Nedajah

DÉDICACE

Je tiens à dédier ce modeste travail à la
lumière de ma vie : mes chers parents , mon
père El Arbi, et ma mère Rabia pour leur plus
grand amour ,

Soutien , encouragement de la patience et de
l'aide contenue pendant mes années d'études .

Ce travail est également dédié :

A mes soeurs douce .

Pour ma cher frère Kemel.

Pour tous mes oncles et tantes

Mes dédicace vont tendrement à mes chère
amies : Ahlem, chaima , Safia ,Youmna ,
Khadija, Tayeb , Ziad, Nedajah

DÉDICACE

Je tiens à dédier ce modeste travail à la
lumière de ma vie : mes chers parents , mon
père Abd elkader, et ma mère Fatima zohra
pour leur plus grand amour ,
Soutien , encouragement de la patience et de
l'aide contenue pendant mes années d'études .
Ce travail est également dédié :
A mes soeurs douce Marwa ,safa .
Pour ma cher frère Abed el hamide.
Pour tous mes oncles et tantes
Mes dédicace vont tendrement à mes chère
amies : Ahlem, Hana, Safia ,Youmna , Khadija,
Tayeb , Ziad, Nedajah

ملخص

الهدف من مشروعنا هو تطوير تطبيق يسمح بالتصنيف تلقائي للمستندات المرقمنة باستخدام تقنية التعرف الضوئي على الحروف ، لتسهيل عملية الأرشفة في كلية العلوم الدقيقة جامعة الشهيد حمة لخضر الوادي.

من بين أمور أخرى ، يتضمن هذا التطبيق عدم تلف المستندات أو فقدانها ، بل ويسمح في الواقع باستخراجها عند الحاجة.

الكلمات المفتاحية: الادارة ، والتصنيف التلقائي ، والمستندات ، والمسح الضوئي ، OCR ، tesseract ، C #.

Résumé

L'objectif de notre projet est le développement d'une application qui permet la classification automatique des documents numérisés en utilisant la technique OCR, pour faciliter le processus d'archivage au niveau de la Faculté des Sciences Exactes, Université ELCHAHID Hamma Lakhder ELOUED.

Entre autre, cette application assure que les documents ne sont pas endommagés ou perdus, et en effet permet de les extraire en cas de besoin.

Mots-clés: administration, classification automatique, documents, numérisation, OCR, tesseract, C#.

Abstract

The aim of our project is the development of an application program that allows the classification of images in a file to facilitate the process of addition, deletion and extraction at the level of archives at the Faculty of Exact Sciences at the University ELCHAHID Hama Lakhder ELOUED.

This program can avoid, damage and loss of documents and easy to extract them when needed.

Keywords: collège, administration, office, document, archive, destination.

Liste des figures

Liste des figures

Figure. I. 1 : Structure de la faculté	7
Figure. I. 2 :Structure de le département d'informatique	8
Figure. I. 3 :Flux d'information.....	8
Figure. I. 4 :Le Enregistrement de document de départ.....	9
Figure. I. 5 :Le Enregistrement de document de arrivée.....	10
Figure. I. 6 :Le document.....	12
Figure .II. 1 L'approche 2TUP	16
Figure .II. 2 : Diagramme de cas de utilisation.....	19
Figure .II. 3: Diagramme de cas de utilisation "Authentification"	19
Figure .II. 4: Diagramme d'activité " Authentification"	20
Figure .II. 5:Diagramme de séquence "Authentification".	21
Figure .II. 6:Diagramme de cas d'utilisation "scanner"	22
Figure .II. 7: Diagramme d'activité "scanner"	22
Figure .II. 8: Diagramme de séquence "scanner"	23
Figure .II. 9:Diagramme de cas d' utilisation "Gérer le document"	24
Figure .II. 10:Diagramme de cas d' utilisation "recherche".	24
Figure .II. 11:Diagramme d'activité "recherche".	25
Figure .II. 12: Diagramme de séquence " recherche "	26
Figure .II. 13: Diagramme de cas d' utilisation " imprimer "	26
Figure .II. 14:Diagramme d' Activité "imprimer".....	27
Figure .II. 15:Diagramme de séquence "imprimer".....	28
Figure .II. 16:Diagramme de cas d' utilisation " supprimer "	28
Figure .II. 17:Diagramme d' Activité " supprimer "	29
Figure .II. 18: Diagramme de séquence " supprimer".....	30
Figure .II. 19: Diagramme de classe	30

Liste des figures

Liste des figures

Figure .III. 1: L' interface principale de l'application.	35
Figure .III. 2: L'interface de l'authentification de l'application.....	35
Figure .III. 3: Interface de scanner.....	36
Figure .III. 4 : présente l'interface après scanner le document.	37
Figure .III. 5: Interface recherche	38
Figure .III. 6: l'interface supprimer ou imprimer le document.	39

Liste des tables

Liste des tables

Tableau. I 1 .synthèse des flux.9

Tableau .II .1:Liste des cas d'utilisateur..... 18

Tableau .II. 2: Liste des classes.....31

SOMMAIRE

Introduction générale.....	1
 Chapitre I : Vision par ordinateur	
1. Introduction.....	4
2. Vision par ordinateur.....	4
2.1. Domaines d'application.....	4
3. Gestion Electronique de Document.....	4
3.1. Classification automatique de document.....	5
a. Numérisation.....	5
b. Scanner	5
c. Algorithmes OCR (Optical character recognition).....	5
• Tesseract.....	6
4. Etude de l'existant	6
4.1. Présentation de la faculté des Sciences Exactes	6
4.1.1. Définition de la faculté.....	6
4.1.2. Structure administrative de la Faculté	6
4.2. Département d'informatique.....	7
❖ Définition de secrétariat	8
4.3. Flux d'information	8
4.4. Archivage	9
4.5. Etude des document.....	10
5. Le critique de l'existant.....	12
5.1. Solution proposée.....	12
6. Conclusion.....	13

Sommaire

Chapitre II: La méthode de développement UML

1. Introduction.....	15
2. Le processus 2TUP.....	15
3. L'UML.....	16
3.1 Définition de l' UML.....	16
3.2. Avantages et inconvénients d'UML.....	17
4. Différent diagramme UML utilisés.....	17
5. Modélisation des besoins.....	18
.1. Identification des acteur	18
.2. Diagramme de cas de utilisation	18
✚ .Cas d'utilisation "Authentification":.....	19
✚ .Diagramme de cas d'utilisation "scanner".....	21
✚ Diagramme de cas d' utilisation " Gérer le document ".....	23
6. Diagramme de classe.....	30
6.1.Description du modèle des classes.....	31
7. Conclusion.....	31

Chapitre III :Développement du système

1. Introduction.....	33
2. Environnement de travail	33
2.1. Langage de programmation "C#".....	33
❖ Les points forts de C#.....	33
2.2. .NET Framework.....	33
2.3. Microsoft Visual studio .Net	34
2.4. SGBD:(système de gestion de base de donnée).....	34
❖ La mise en œuvre de instances et des bases de données SQL Server	34
3. Présentation d'application	34
3.1. Interface de « démarrage »	35
3.2. Interface « d'authentification »	35
3.3. Interface «scanner ».....	36
3.4. Interface « recherche »	38
3.5. Interface « imprimer et supprimer».....	39

Sommaire

4. Conclusion.....	40
Conclusion générale.....	41
Références	42

Introduction générale

La vision par ordinateur permet d'imiter le système visuel humain (association entre le sens de la vue et un ensemble de processus cognitifs). elle vise à traiter des données visuelles acquises via des capteurs visuels (par exemple : caméras), et extraire des informations pertinentes de manière à percevoir, analyser et comprendre la scène observée par le biais de modèles fondés sur la géométrie, la physique, la biologie, les statistiques et la théorie d'apprentissage.

Aujourd'hui, avec le développement de la gestion électronique des documents en termes de programmes, l'utilisation de ces documents passe de l'ancien au moderne: la numérisation, l'édition, la classification, l'indexation et l'archivage.

En fait, la gouvernance électronique est devenue plus précieuse que jamais, ainsi que la capacité de répondre à un large éventail de besoins, en plus du développement continu, ce qui en fait un outil encore plus important.

Dans ce contexte, fournit aux utilisateurs certaines options. Cependant, il semble faire face à certaines de ces opérations et de l'utiliser très difficile pour la plupart des gestionnaires parce qu'ils ont besoin de structures de main lourdes et procédures.

Problème que nous offrons, comment exploiter la vision par ordinateur dans le domaine de la gestion des documents électroniques et classés automatiquement, notre objectif est de, dans le cadre de ce projet, la création d'une application au profit de administrations, établissements et les entreprises afin d'assurer la fiabilité et la maniabilité simple et flexibilité permet également aux utilisateurs de se concentrer davantage sur leurs tâches tout en réduisant le risque d'oublier et l'erreur d'améliorer considérablement et donc les procédures traditionnelles en utilisant la technologie de reconnaissance optique des caractères.

Le présent mémoire est structuré en trois chapitres comme suit:

- ✓ Le premier chapitre présente une vue générale de la vision par ordinateur etnotamment la problématique de reconnaissance de caractères et la gestion électronique de document.
- ✓ .Le deuxième chapitre concerne la conception de notre application à l'aide des diagramme UML selon la démarche de processus 2TUP.
- ✓ Le dernier chapitre présenté les outils utilisé lors de la réalisation de notre application , aussi bien quelques interface représentant notre application.
Enfin, nous conclurons ce projet par une conclusion générale.

Chapitre I

Vision par ordinateur

1. Introduction

L'objectif de cette étude est de concevoir de réaliser un système d'information pour conception et développement application de classification des images de documents ,sur le niveau de la faculté des sciences exactes à l' université ELCHAHIDE HammaLakhder d'ELOUED .

La vision par ordinateur représente la technologie centrale de l'analyse d'image automatique, qui est utilisée dans de nombreux axes de recherche. Dans ce chapitre nous allons explorer le domaine de la vision par ordinateur et notamment la problématique de allons explorer connaissance de caractères.

2. Vision par ordinateur

La vision par ordinateur est une branche de l'intelligence artificielle dont le principal but est de permettre à une machine d'analyser, traiter et comprendre une ou plusieurs images prises par un système d'acquisition semblable à la perception visuelle. En outre, est une approche consiste à tenter d'imiter la vision humaine ou animale via des composants électroniques. En fait elle vise à traiter des données visuelles acquises via des capteurs visuels (par exemple : caméras), par le biais de modèles fondés sur la géométrie, la physique, la biologie, les statistiques et la théorie d'apprentissage. La vision par ordinateur a aussi été décrite comme une initiative dans l'automatisation et l'intégration d'une vaste gamme de technologies dans certains axes de recherche.[1]

2.1. Domaines d'application

En raison des capacités étendues de l'ordinateur dans divers domaines, de nombreuses applications basées sur vision par ordinateur ont été mentionnées dans certains domaines, parmi eux: domaines commercial et économique, domaines de la science et de la recherche, ingénierie et domaines scientifiques et spatiaux... etc. En outre, la technologie OCR incorporée dans la vision par ordinateur est utilisée dans plusieurs applications afin de manipuler et modifier le texte lu contenant dans une image ou un document numérisé.[2]

3. Gestion Electronique de Document

La gestion électronique des documents (GED ou en anglais DMS pour Document Management System ou EDM1 pour Electronique Document Management) désigne un procédé informatisé visant à organiser et gérer des informations et des documents électroniques au sein d'une organisation. Le terme GED désigne également les logiciels permettant la gestion de ces contenus documentaires.

La GED met principalement en œuvre des systèmes d'acquisition (exemple d'utilisation: la numérisation de masse de documents papiers), d'indexation, de classement, de stockage d'information, d'accès (navigation et recherche) et de diffusion des documents. La GED participe ainsi aux processus de collaboration, de capitalisation et d'échange d'informations. Elle prend en compte le besoin de gestion des documents selon leur cycle de vie, de la création à l'archivage en passant par la gestion des différentes versions.

Les solutions et projets de GED peuvent permettre des gains en qualité et en coût rapide pour les organisations, qu'il s'agisse d'entreprises privées ou d'administrations. À titre d'exemple, le retour sur investissement d'un projet « standard » de dématérialisation de factures est inférieur à un an.

Une GED dispose de fonctions de classification / navigation et également d'un moteur de recherche qui permettent de retrouver les contenus gérés, au moyen de vues, de recherches structurées ou plein texte (full texte).[3]

On va détailler par la suite la partie concernant la classification automatique :

3.1. Classification automatique de document

Il existe des étapes pour la classification des documents sont :

a. Numérisation

est une forme d'authentification électronique, où le transfert de documents de processus de numérisation Sur un support électronique, et prendre deux formes de numérisation sous la forme d'images, la numérisation du format de texte, où peut Introduire quelques conversions et modifications, après avoir traité le texte à l'aide d'un programme spécial Pour la reconnaissance de caractères.

b. Scanner

(scanner de langue anglaise) d'origine du scanner est dispositif de lecture et d'entrée d'imagerie des documents imprimés et de les convertir en un signal numérique pour entrer dans l'ordinateur en tant que retouche photo plat pour faire ressortir l'origine possible, et les scanners dans les années soixante du XIXe siècle, a été créé un scanner que nous connaissons aujourd'hui En 1957 par Kirsch le contact général aux États-Unis, le Bureau national des statistiques, où le premier scanner taille de l'image 5 x 5 cm, exactement 176 pixels.[4]

c. Algorithmes OCR (Optical character recognition)

Le mot OCR (en anglais : Optical character recognition) signifie reconnaissance optique de caractères ou reconnaissance de texte, une technologie qui vous permet de convertir différents types de documents tels que les documents papiers scannés, les fichiers

PDF ou les photos numériques en fichier modifiables et interrogeables. Supposez que vous ayez un document au format papier, par exemple un article de journal, une brochure ou un contrat, que votre partenaire vous a envoyé en pièce jointe au format PDF. Un scanner ne suffit pas pour extraire les informations pertinentes de ces documents pour les retranscrire par exemple dans le format éditable Microsoft Word. Tout ce qu'un scanner peut faire c'est de réaliser une image du document. Et celle-ci n'est autre qu'une collection de points noirs et blancs ou de couleur, ordonnés sous forme de grille, et connue en langage technique sous le terme de « trame ». Pour lire et exploiter les informations issues de documents scannés, d'images numériques ou de PDF d'images seulement, il faut un logiciel OCR qui reconnaît dans les images des lettres, reconstitue des mots, puis des phrases entières. Le logiciel vous permet ainsi d'avoir accès au contenu proprement dit des documents que vous pouvez ensuite retravailler.[5]

- **Tesseract**

Tesseract est un moteur de reconnaissance de texte open source (OCR) , disponible sous licence apache 2.0.II peut être utilisé directement, ou (pour les programmeurs)en utilisant une API pour extraire de texte imprimé à partir d'image .Il prend en charge une grande variété de langues .

Tesseract n'a pas d'interface graphique intégrée, mais plusieurs sont disponible sur la page 3rd Party. [6]

4. Etude de l'existant

4.1.Présentation de la faculté des Sciences Exactes

Dans cette section, nous fournissons des informations sur la faculté

4.1.1. Définition de la faculté

Elle est ressortie de la restructuration de l'université en juillet ici à 2015 après avoir fait partie de la faculté des sciences et de la technologie , et comprend les sections suivante :

- ✓ Département de mathématique
- ✓ Département d'informatique
- ✓ Département de chimie
- ✓ Département de physique

4.1.2. Structure administrative de la Faculté

La directeur de faculté est divisée en 5 service

-Secrétariat général de la faculté .

-Vice rectorat du doyen chargé des études et les question liées aux étudiants .

- Vice rectorat de la formation supérieure de Post Graduation de l'habilitation Universitaire la Recherche Scientifique (PGRS).
 - Bibliothèque
 - section (département d' informatique ,département de mathématique ,département chimie ,département de physique).
- Et construire la division de administrative de la faculté des sciences exactes dans la figure.

I.1 [7]

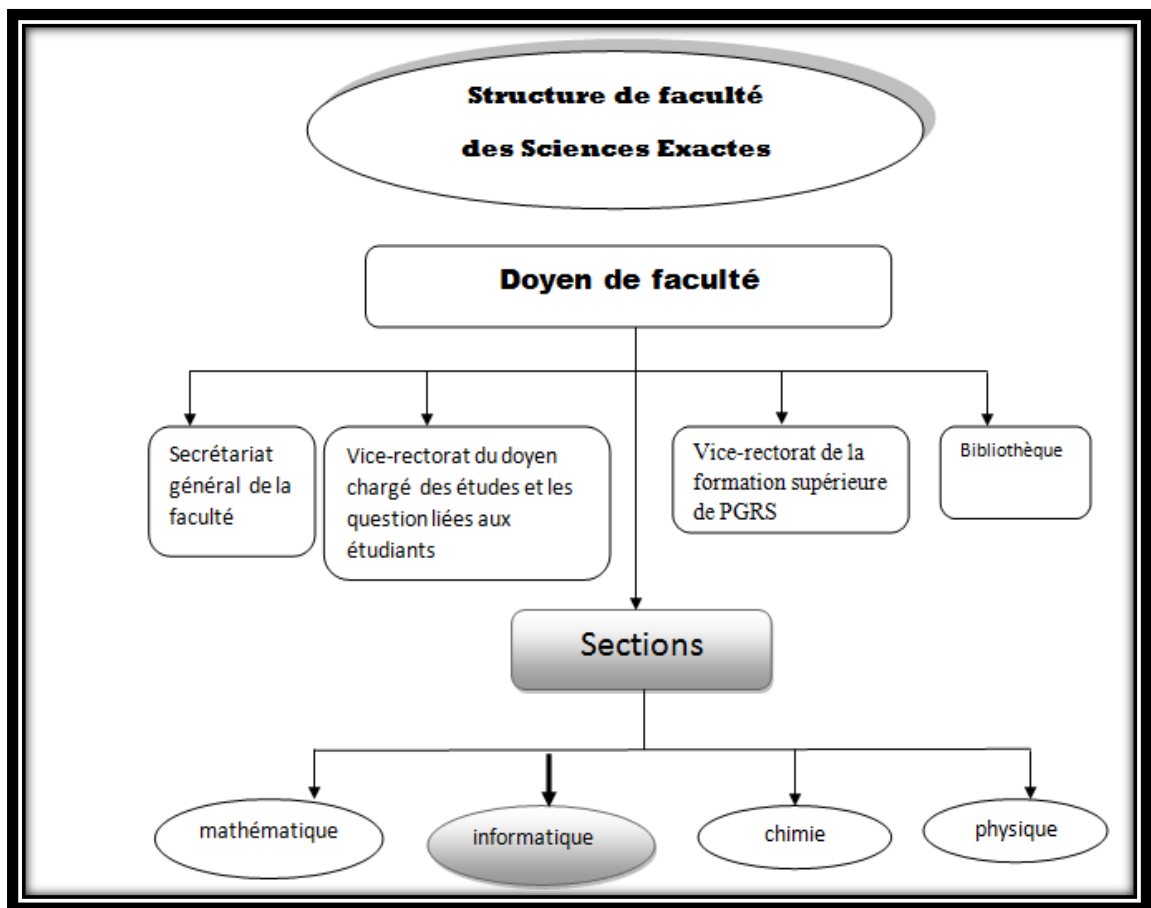


Figure. I.1:Structure de la faculté

4.2. Département d'informatique

C'est une division de la Faculté de Sciences Exactes, Les personnel administratif et le personnel enseignant de cette section surveilleront et suivront les étudiants au niveau du baccalauréat et du master du nouveau système LMD. [7]

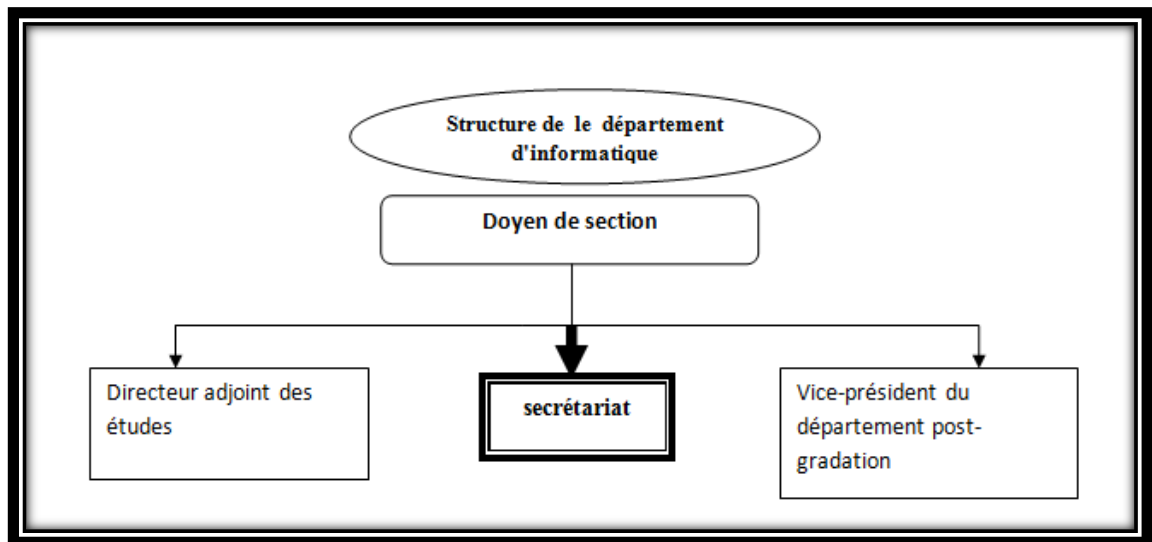


Figure. I. 2: Structure de le département d'informatique

❖ Définition de secrétariat

C'est une partie essentielle de l'administration et fournit de nombreux services, parmi lesquels l'enregistrement des documents départ, arrivée et stockés manuellement.

4.3. Flux d'information

Le diagramme de flux des informations permet de montrer les différents acteurs internes et externes, ainsi que les différent flux qui circulent entre ces acteurs. La figure 03 présente les flux d'informations dans notre système.

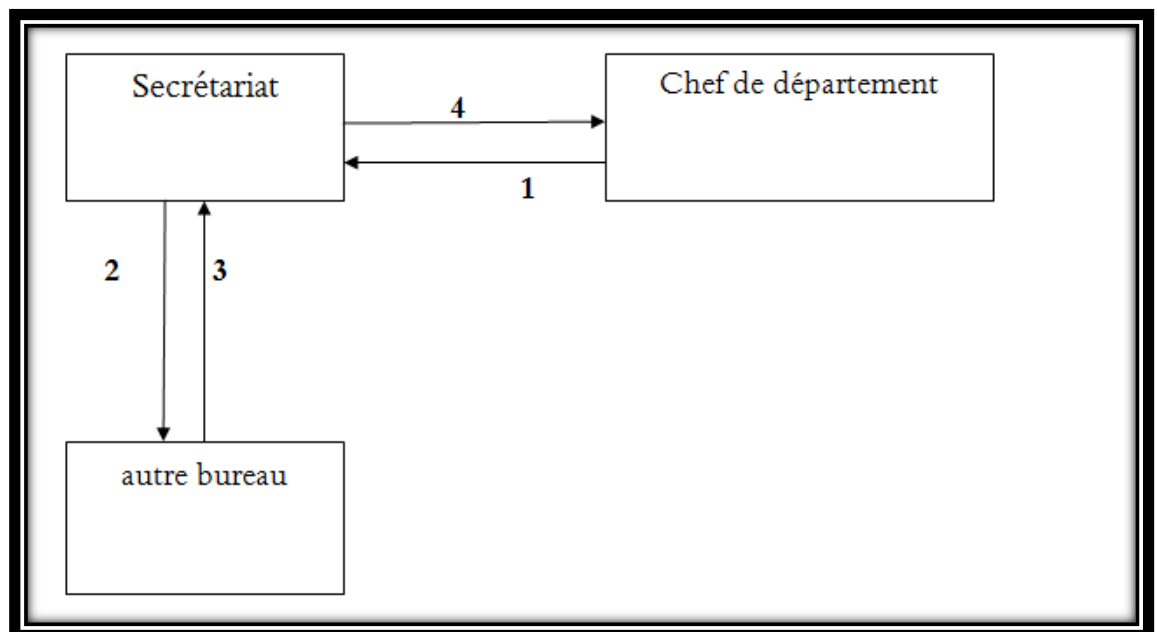


Figure. I. 3: Flux d'information.

La table 01 représente la synthèse des flux.

Table. I.1: synthèse des flux.

Flux	Désignation
1	Demande un document
2	Envoyer un document
3	Reçu un document
4	La réponse

4.4. Archivage

Le processus de stockage de données et d'informations sur un support électronique pendant une longue période de temps est important et rare, pour des raisons historiques et de sécurité, et récupéré si nécessaire.[8]

✓ Enregistrement de document de départ

Les départs sont consignés dans un registre délivré et enregistré à une date, un numéro de série et le thème du document

Et construire tableau de l'enregistrement de document de départ dans le Figure I. 3.

DEPART			الصادر		
تاريخ و رقم الرسالة DATE ET N° DE LA CORRESPONDANCE	عدد الوثائق NOMBRE DE PIECES	المرسل إليه DESTINATAIRE	الموضوع OBJET	رقم الأرشيف N° ARCHIVES	الملاحظات OBSERVATIONS
Le..... في N°					
Le..... في N°					
Le..... في N°					
Le..... في N°					
Le..... في N°					
Le..... في N°					
Le..... في N°					

Figure .I. 4. Le Enregistrement de document de départ

✓ Enregistrement de document de arrivée

Les arrivées sont enregistrées dans un enregistrement appelé entrant et enregistré par date, numéro de série et sujet du document.

Et construire tableau de le Enregistrement de document de entrée dans le Figure. I.4.

ARRIVEE				الوارد
تاريخ الوصول Date d'Arrivee	تاريخ ورقم الرسالة Date et N° de Correspondance	الباعث Emetteur	الموضوع Objet	تاريخ ورقم الاجابة Date et N° de La réponse
Le في N° رقم	du في N° رقم			Le في N° رقم
Le في N° رقم	du في N° رقم			Le في N° رقم
Le في N° رقم	du في N° رقم			Le في N° رقم
Le في N° رقم	du في N° رقم			Le في N° رقم
Le في N° رقم	du في N° رقم			Le في N° رقم
Le في N° رقم	du في N° رقم			Le في N° رقم

Figure .I. 5:Le Enregistrement de document de arrivée.

4.5. Etude des document

Durant l'étape de collecte d'information ; on a recensé les différents fichiers qui interviennent dans les procédures de travail dans le but de recensement les information nécessaires pour l'élaboration du dictionnaire de données .

En effet , une étude détaillée de chacun de ces fichiers s'avère plus que substantielle et celle-ci se fera en trois parties:

- Caractérisation du fichiers .
- Acheminement.
- Analyse du fichiers.

Les types utilisées pour les rubriques des fichiers sont:

C : Caractère.

AN : Alphabétique Numérique .

D : Date.

Les observation éventuelles:

JJ/MM/AAAA : Jour/mois/année.

De.....a..... : De Jour/mois/année a. Jour/mois/année.

Pour le courrier sortant et entrant nous avons plusieurs formes d'écriture de document:

	Date	Numéro
Document	Ordinateur	Ordinateur
Document	Ordinateur	Manuel
Document	Dateur	Manuel

A partir de les modèles précédents on va parler sur le premier modèle

***Ordinateur -Ordinateur ***

La figure 05 représente un modèle de document.

الجمهورية الجزائرية الديمقراطية الشعبية
République Algérienne Démocratique et Populaire
وزارة التعليم العالي والبحث العلمي
Ministère de l'Enseignement Supérieure et de la Recherche Scientifique
جامعة الشهيد حمه لخضر - الوادي
Université Echahid Hamma Lakhder- El Oued

كلية العلوم الدقيقة
نيابة العمادة لما بعد التدرج
والبحث العلمي والعلاقات

Faculté des Sciences Exacte
Le Vice-Doyen Chargé de la Post-
Graduation, de la Recherche
Scientifique et des Relation Extérieures

الوادي في: 2018/02/26

الرقم: 2018/VCPRSR/039

الموضوع : ب/خ تسوية ملفات التسجيل لطلبة الدكتوراه

Figure .I. 6:Le document.

5. Le critique de l'existant

-Les documents sont plus susceptibles d'être endommagés et perdus à la suite d'une action aléatoire.

-Manque difficulté à chercher e de logiciel pour organiser les opérations d'archivage.

-Difficulté à chercher.

5.1. Solution proposée

Nous allons créer application de classification a partir de quelques fonctionnalités pour automatiser le procédure de classification par l'utilisation de technologie OCR afin de répondre à tous les besoins des utilisateurs.

6. Conclusion

Dans ce chapitre, nous avons parlé de la vision par ordinateur et la technologie de reconnaissance optique de caractères et ses domaines d'applications.

Dans ce qui suit nous allons entamer la première phase de la conception et de modélisation de notre projet, afin d'identifier les différentes fonctionnalités de l'application.

Chapitre II

**La méthode de développement
UML**

1. Introduction

Une méthode de développement définit à la fois un langage de modélisation et la démarche (processus) à suivre lors de la modélisation.

Dans ce chapitre nous présentons un outil de modélisation UML et le processus de développement logiciel, tel que le développement avec l'UML demande un processus qui permet d'envisager la marche à suivre lors de la conception.

UML est conçu pour être indépendant du processus de développement et le choix de celui-ci dépend le type de projet, pour notre système on choisit le processus 2TUP, car il traite le projet du domaine de système d'information.

2. Le processus 2TUP

Le processus 2TUP (Two Track Unified Process) est un processus unifié qui a pour but d'apporter une réponse aux contraintes de changement fonctionnelles et techniques qui s'imposent aux systèmes d'information. 2TUP propose un cycle de développement qui dissocie les aspects techniques des aspects fonctionnels. Il part du constat que toute évolution imposée au système d'information peut se décomposer et se traiter parallèlement, suivant un axe fonctionnel et un axe technique. Il distingue ainsi deux branches (fonctionnelle et technique) dont les résultats sont fusionnés pour réaliser le système. La figure. II.1 illustre les principales étapes du processus de développement.[9]

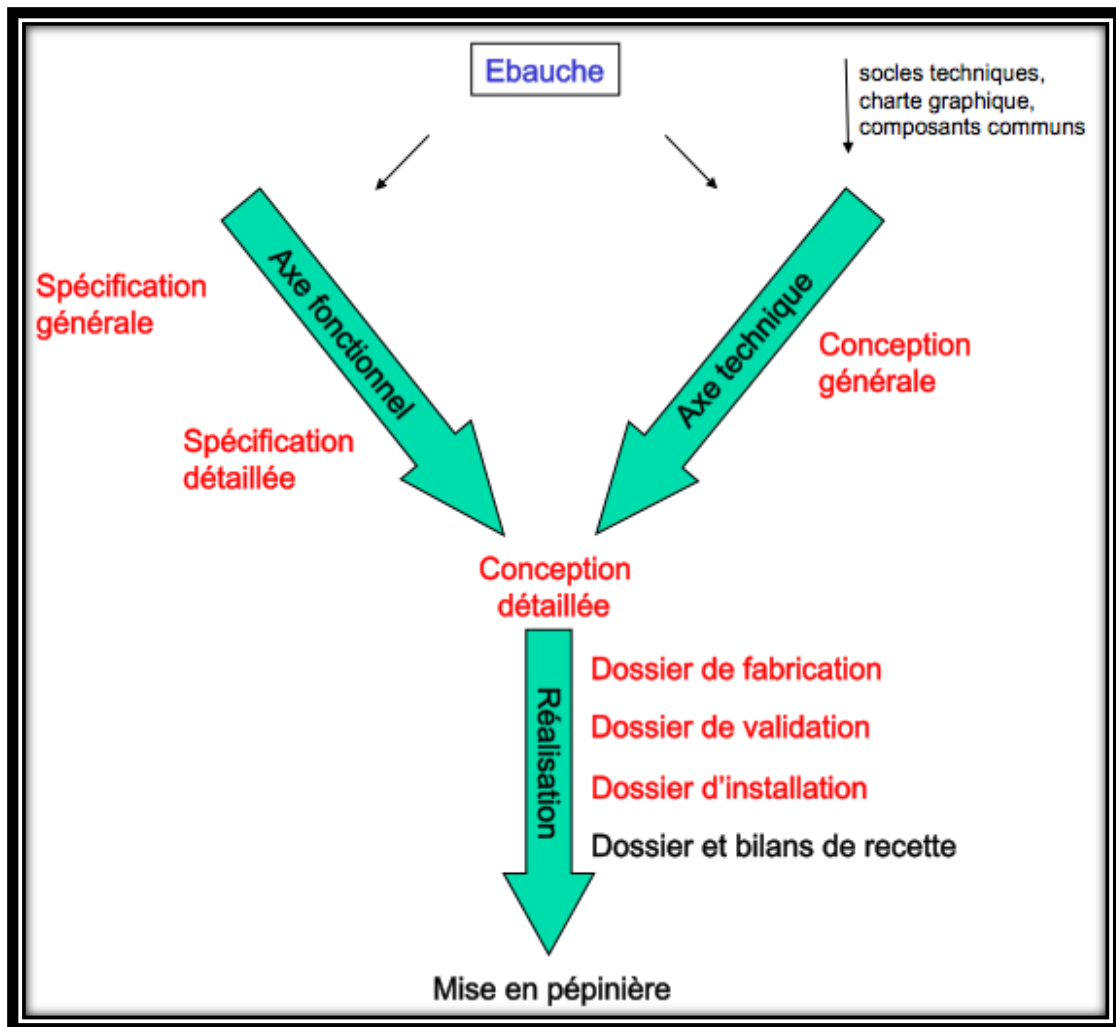


Figure .II.1: L'approche 2TUP [9].

3. L'UML

3.1. Définition de l' UML

L'UML se définit comme un langage de modélisation graphique et textuel.

Il est destiné à comprendre et décrire des besoins, spécifier et documenter les systèmes, et sert aussi à esquisser des architectures logicielles, concevoir des solutions et communiquer des points de vue.

L'UML unifie à la fois les notations et les concepts orientés objet. Il ne s'agit pas d'une simple notations et graphique, car les concepts transmis par un diagramme ont une sémantique précise et sont porteurs de sens au même titre que les mots d'un langage.[10]

3.2. Avantages et inconvénients d'UML

❖ **Avantage d' UML**

UML est un langage normalisé:

- gage de stabilité
- gain de précision
- encourage l'utilisation d' outils

UML est support de communication performants:

- il cadre l'analyse.
- il facilite la compréhension de représentation abstraites complexes.
- son caractère polyvalent et sa souplesse en font un langage universel.

❖ **Inconvénient d' UML**

- La mise en pratique d'UML nécessite un apprentissage et par une période d' adaptation.
- Le processus (non couvert par UML) est une autre clé de la réussite d' un projet.[11]

4. Différent diagramme UML utilisés

L'UML offre beaucoup de diagrammes qui servent à la modélisation des système, nous allons présenter la définition de quelque diagramme :

Diagramme de cas d' utilisation

Le diagramme de cas d'utilisation représente la structure des grandes fonctionnalités nécessaire aux utilisateurs du système , C'est le premier diagramme du modèle UML, celui qui assure la relation entre l'utilisateur et les objets que le système met œuvre.

Diagramme de séquence

Le diagramme de séquence représente la succession chronologique des opérations réalisées par un acteur . Il indique les objets que l'acteur va manipuler et les opérations qui font passer d'un objet à l'autre.

Diagramme de classes

Le diagramme de classes est généralement considéré comme le plus important dans un développement orienté objet. Il représente l'architecture conceptuelle du système : il décrit les classes que le système utilise, ainsi que ceux-ci représentent un emboîtement conceptuel (héritage) ou une relation organique (agrégation).

Diagramme d'activité

Les diagrammes d'activités permettent de mettre l'accent sur les traitements. Ils sont donc particulièrement adaptés à la modélisation du cheminement de flots de contrôle et de flots de données. Ils permettent ainsi de représenter graphiquement le déroulement d'un cas d'utilisation.

5. Modélisation des besoins

La phase d'analyse a pour objectif d'identifier les acteurs qui interagissent avec le système et spécifier les fonctionnalités du système. Cette phase sert aussi à déterminer pour acteur ses droits d'accès à l'utilisation du système. [12]

5.1. Identification de acteur

La première étape de modélisation consiste à définir le périmètre du système, le contour de l'organisation. La seconde sert à modéliser et identifier les entités qui interagissent avec le système qui sont les acteurs.

Définition: Un acteur représente rôle joué par une entité externe (utilisateur humain, dispositif matériel ou autre système) qui interagit directement avec le système étudié.

 Utilisation.

L'acteur (Utilisation): Reçoit et envoyer les documents, Scanner, Enregistrer.

5.2. Diagramme de cas d'utilisation

Tableau .II.1 :Liste des cas d'utilisateur.

Cas d'utilisation	Acteur
Scanner	Secrétariat
Gérer de document	Secrétariat

Dans la figure suivant nous présentons notre diagramme de cas d'utilisation

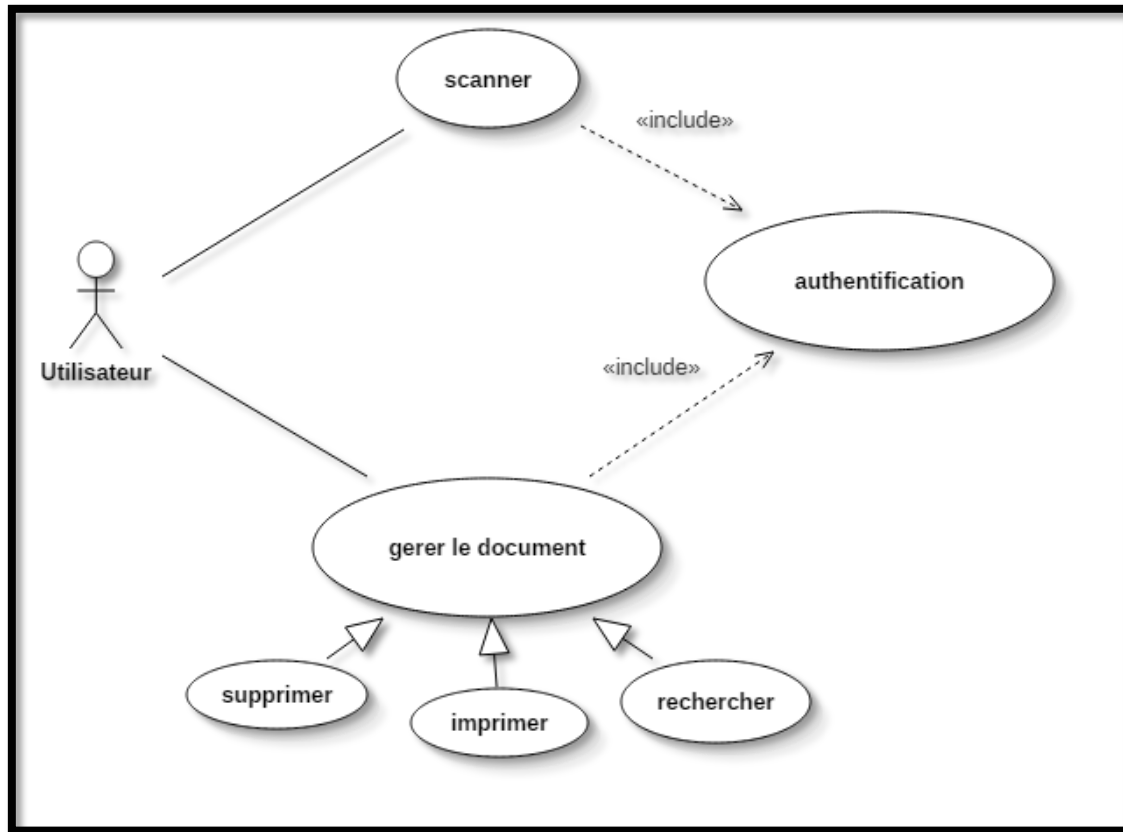


Figure. II.2: Diagramme de cas de utilisation.

✚ Cas d'utilisation "Authentification":

But : authentification et autorisation d' accès.

Action :

- saisir le nom et mot de passe.
- valider puis connecter.

a. Diagramme de cas d'utilisation:

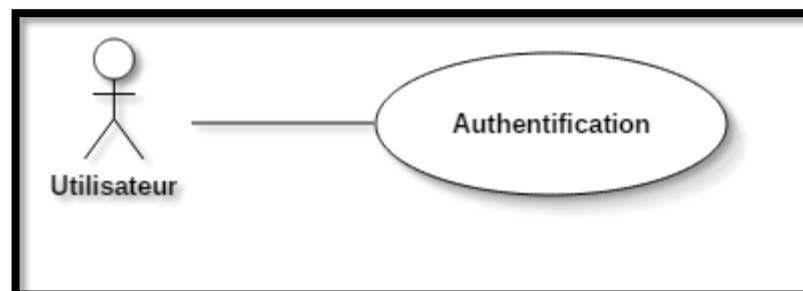


Figure. II .3: Diagramme de cas d'utilisation "Authentification".

b. Diagramme d'activité

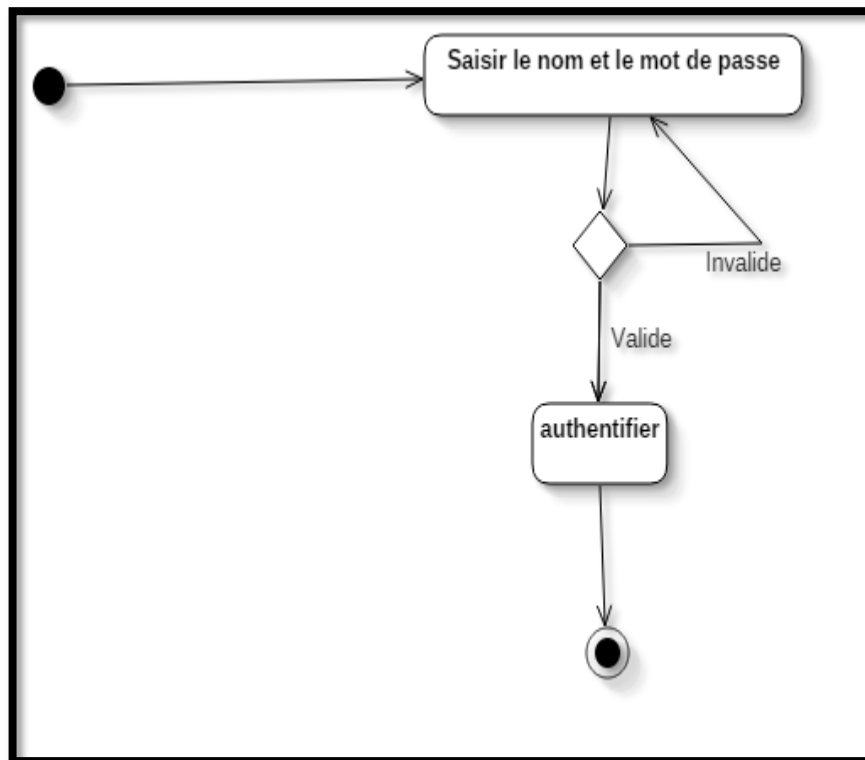


Figure .II. 4: Diagramme d'activité " Authentification".

c . Diagramme de séquence

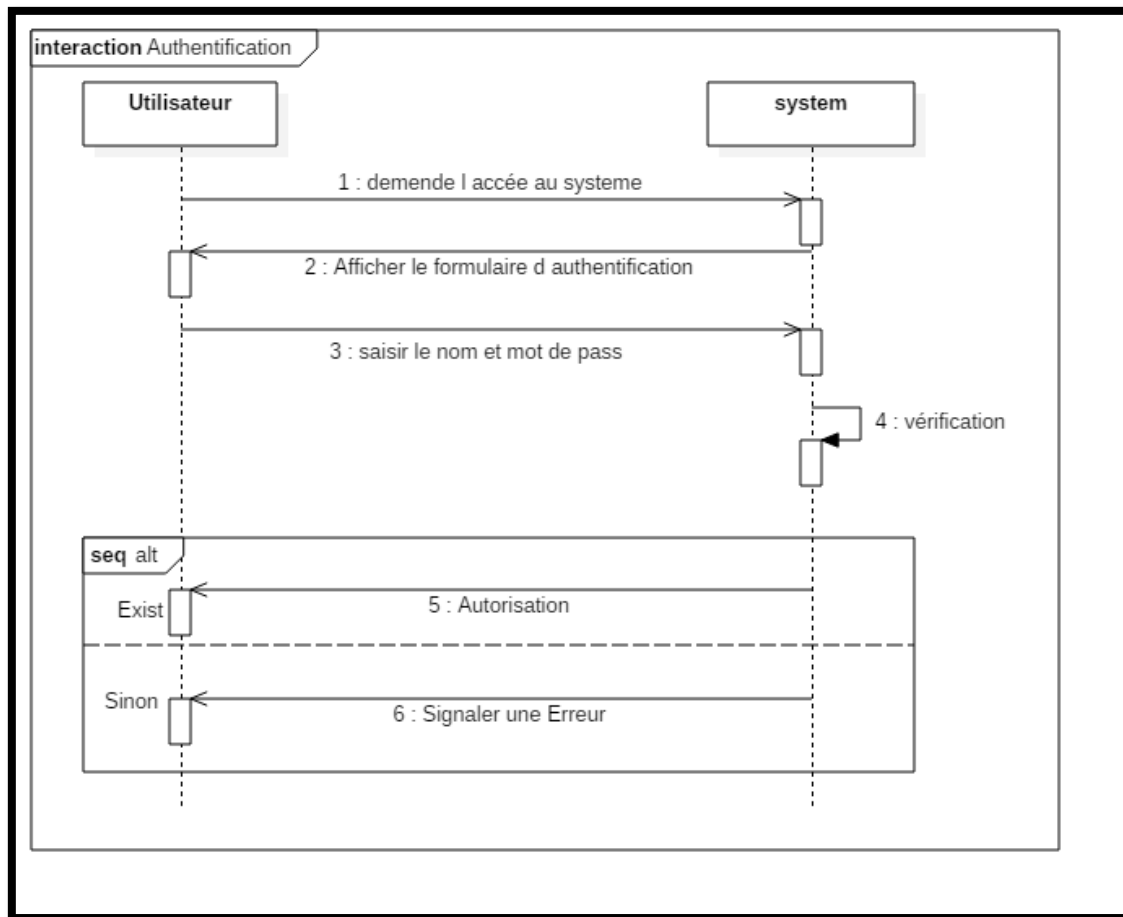


Figure .II. 5:Diagramme de séquence "Authentification".

✚ **Diagramme de cas d'utilisation "scanner"**

But : Transformer le document à une image.

Action :

1. Après l'Authentification
2. scannée le document.

a. Diagramme de cas d'utilisation



Figure .II. 6:Diagramme de cas d'utilisation "scanner"

b. Diagramme d' activité

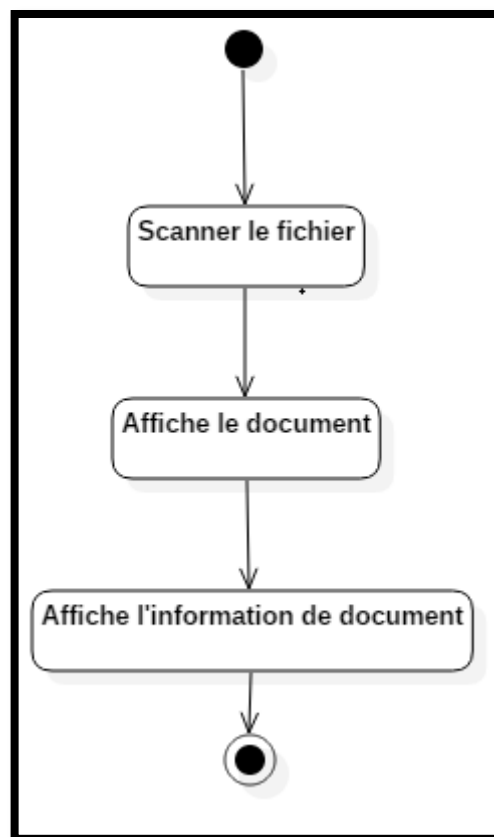


Figure. II .7: Diagramme d'activité "scanner".

c. Diagramme de séquence

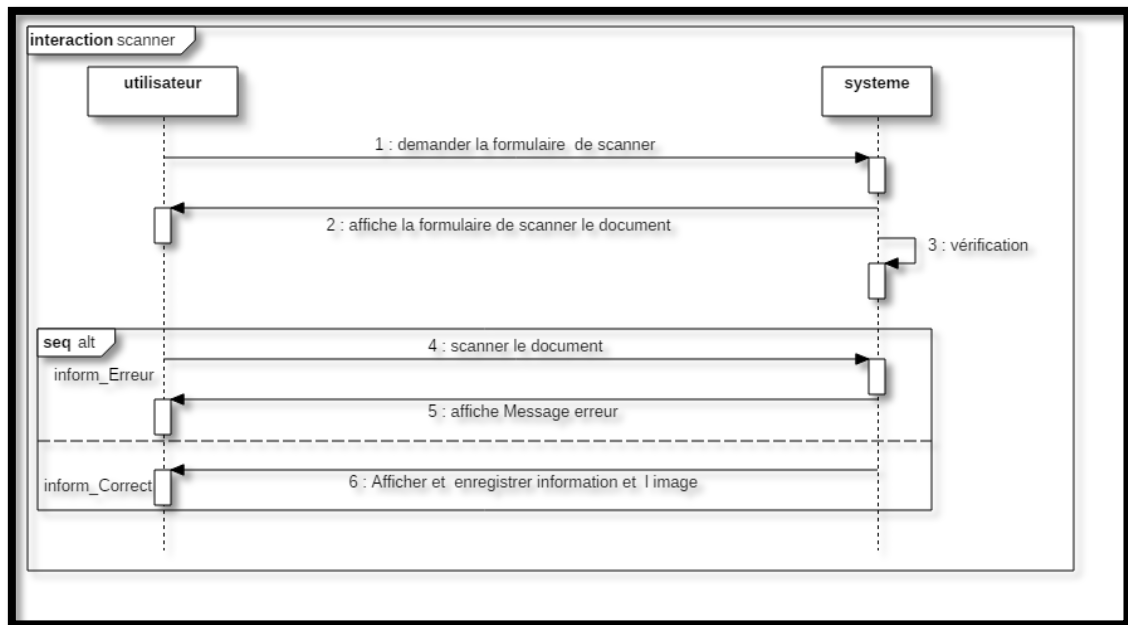


Figure .II.8: Diagramme de séquence "scanner".

Diagramme de cas d'utilisation " Gérer le document "

But: une vérification et une validation pour faire l'action de "recherche" et après les différents action "Modifier , Supprimer et Imprimer"

Action:

- après l' authentification.
- Rechercher un document.
- Modifier , supprimer ou imprimer un document.

✓ Diagramme de cas d'utilisation:

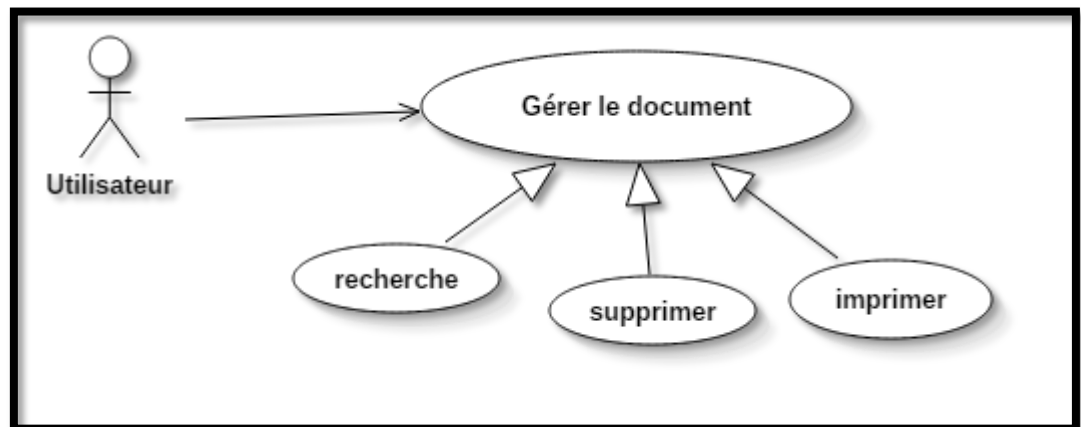


Figure .II.9:Diagramme de cas d' utilisation "Gérer le document"

a. Diagramme de cas d'utilisation "recherche"

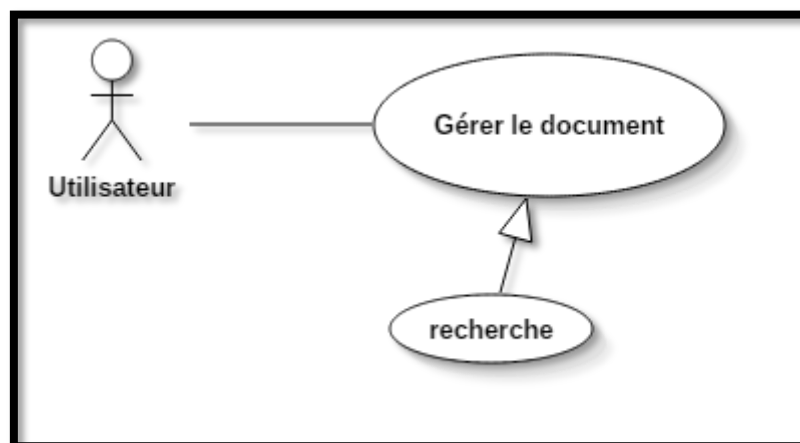


Figure .II.10:Diagramme de cas d'utilisation "recherche".

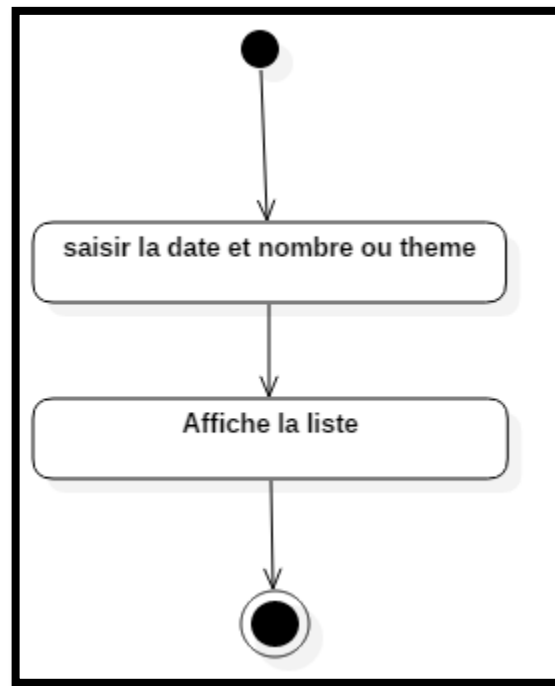
b. Diagramme d'activité "recherche"

Figure .II.11:Diagramme d'activité "recherche".

c. Diagramme de séquence "recherche"

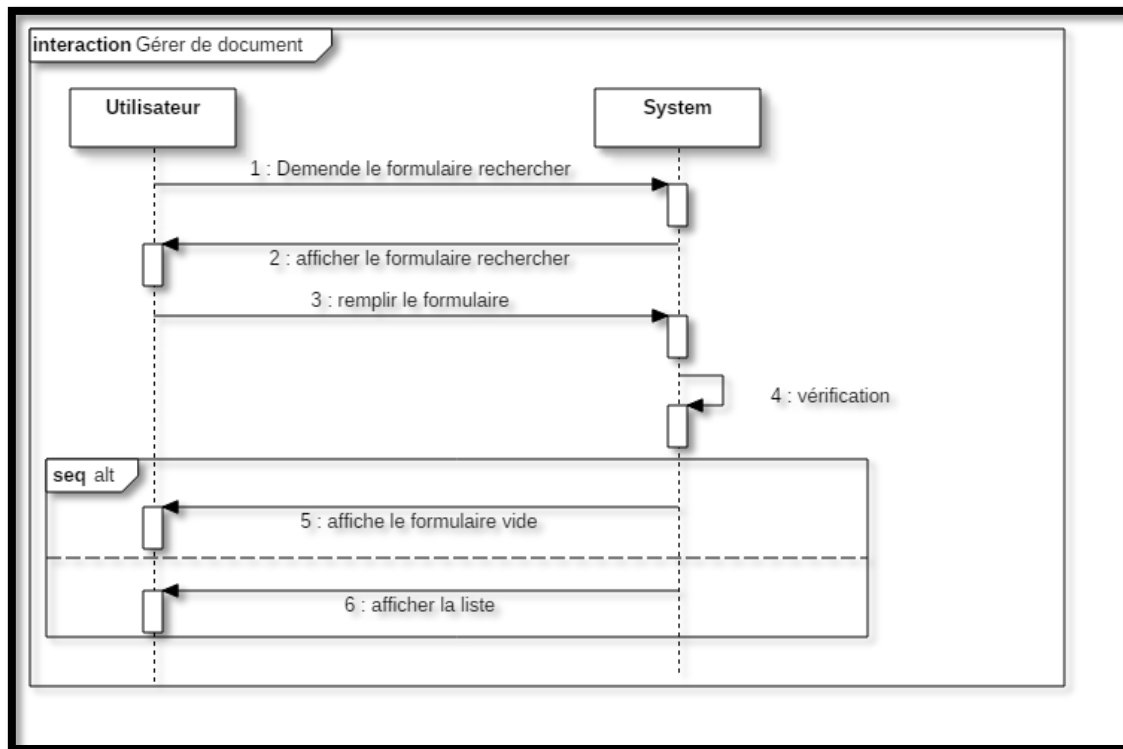


Figure .II.12: Diagramme de séquence " recherche ".

a. Diagramme de cas d' utilisation "imprimer "

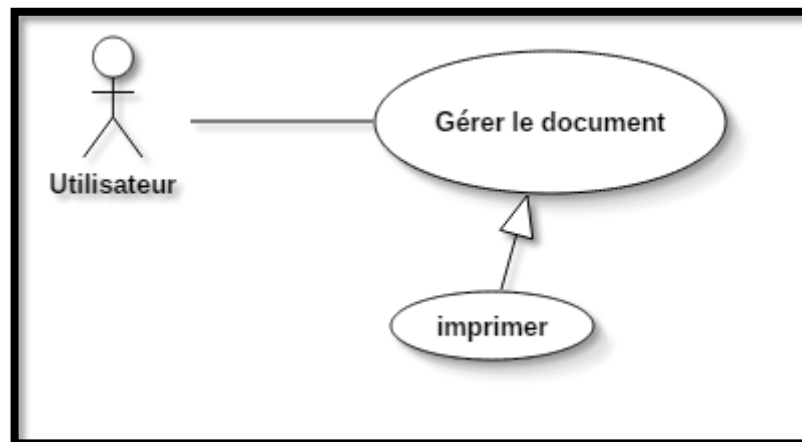


Figure .II.13: Diagramme de cas d' utilisation " imprimer " .

b. Diagramme d'activité "imprimer"

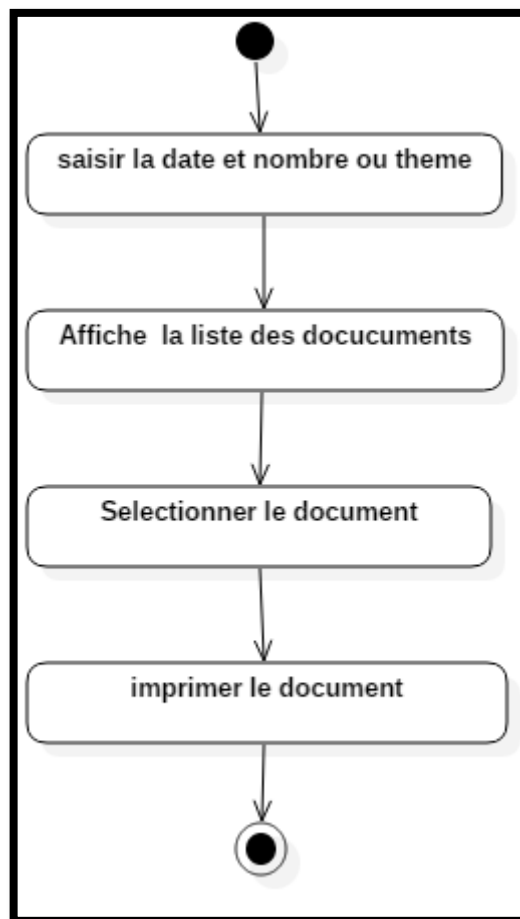


Figure .II.14:Diagramme d'activité "imprimer".

c. Diagramme de séquence "imprimer"

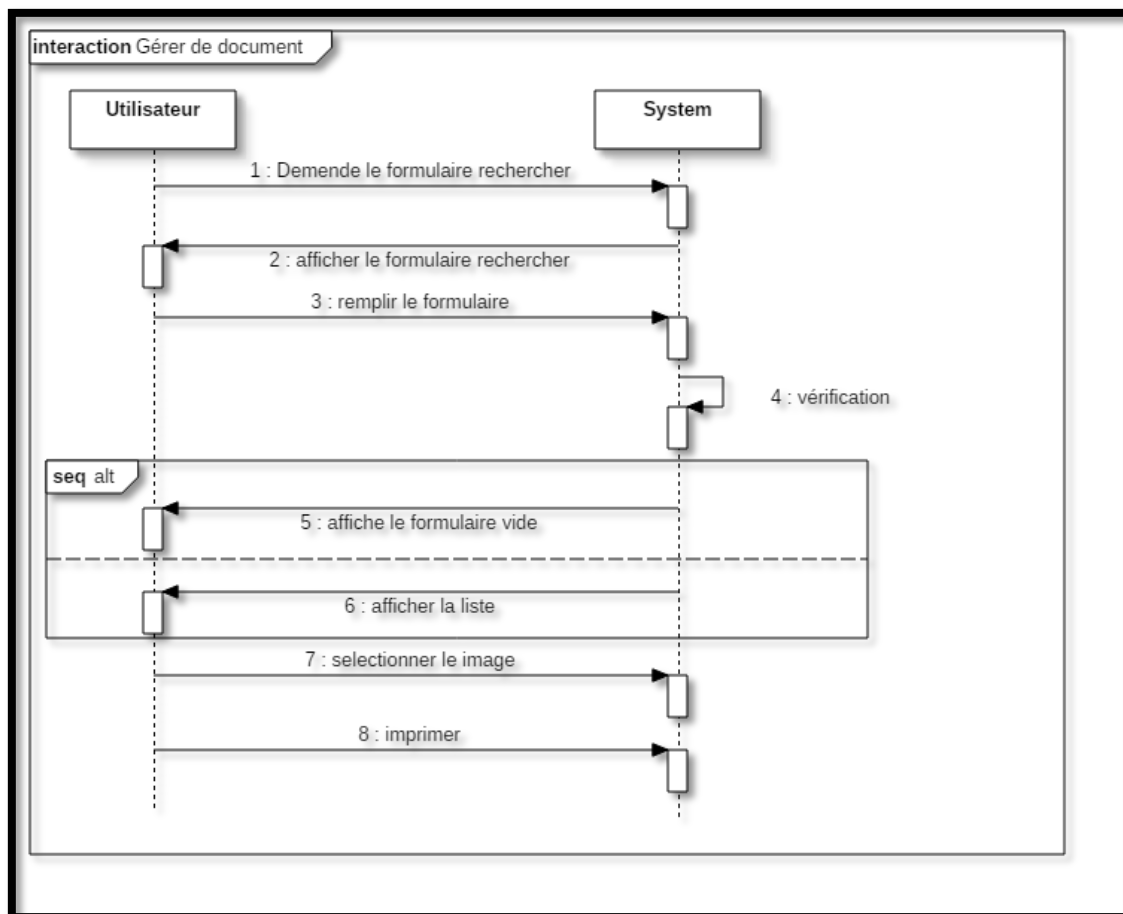


Figure .II.15:Diagramme de séquence "imprimer".

a. Diagramme de cas d' utilisation " supprimer "

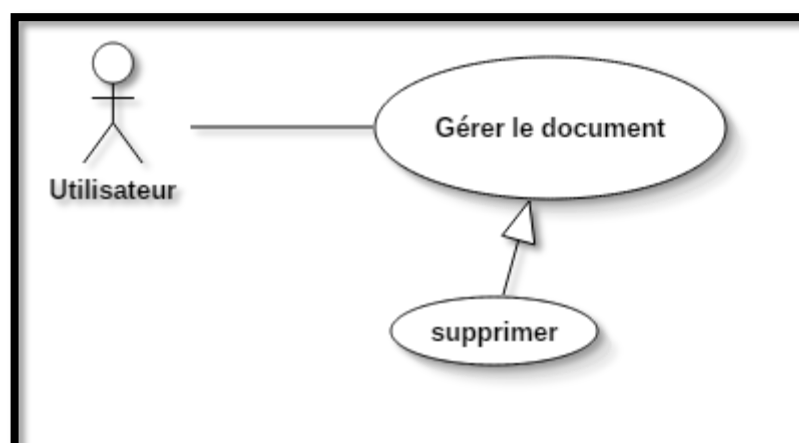


Figure .II.16:Diagramme de cas d' utilisation " supprimer ".

b. Diagramme d'activité " supprimer "

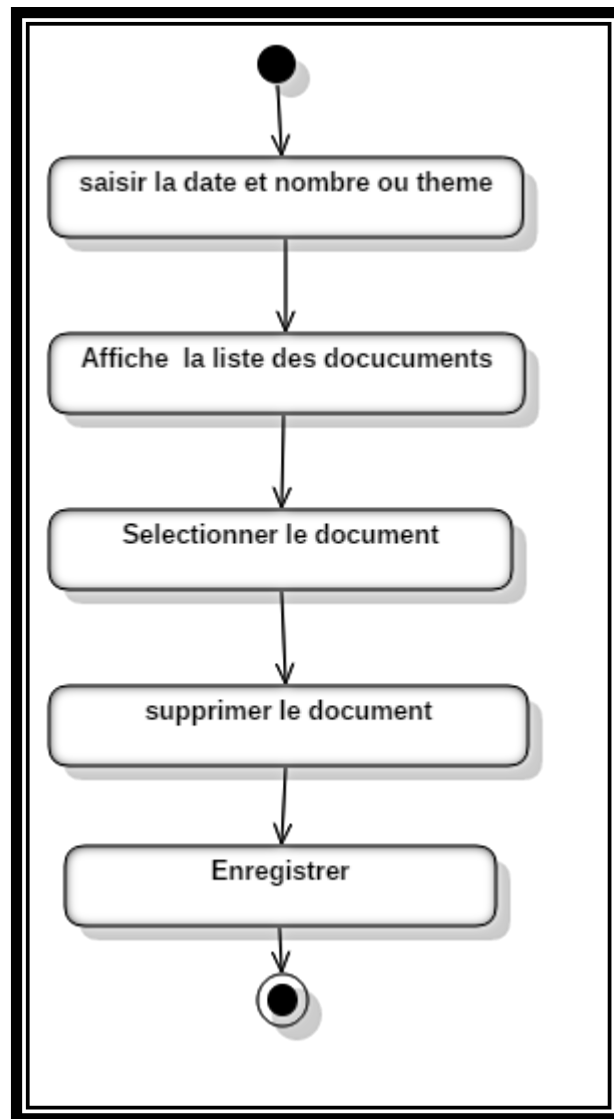


Figure .II.17: Diagramme d'activité " supprimer ".

c. Diagramme de séquence " supprimer "

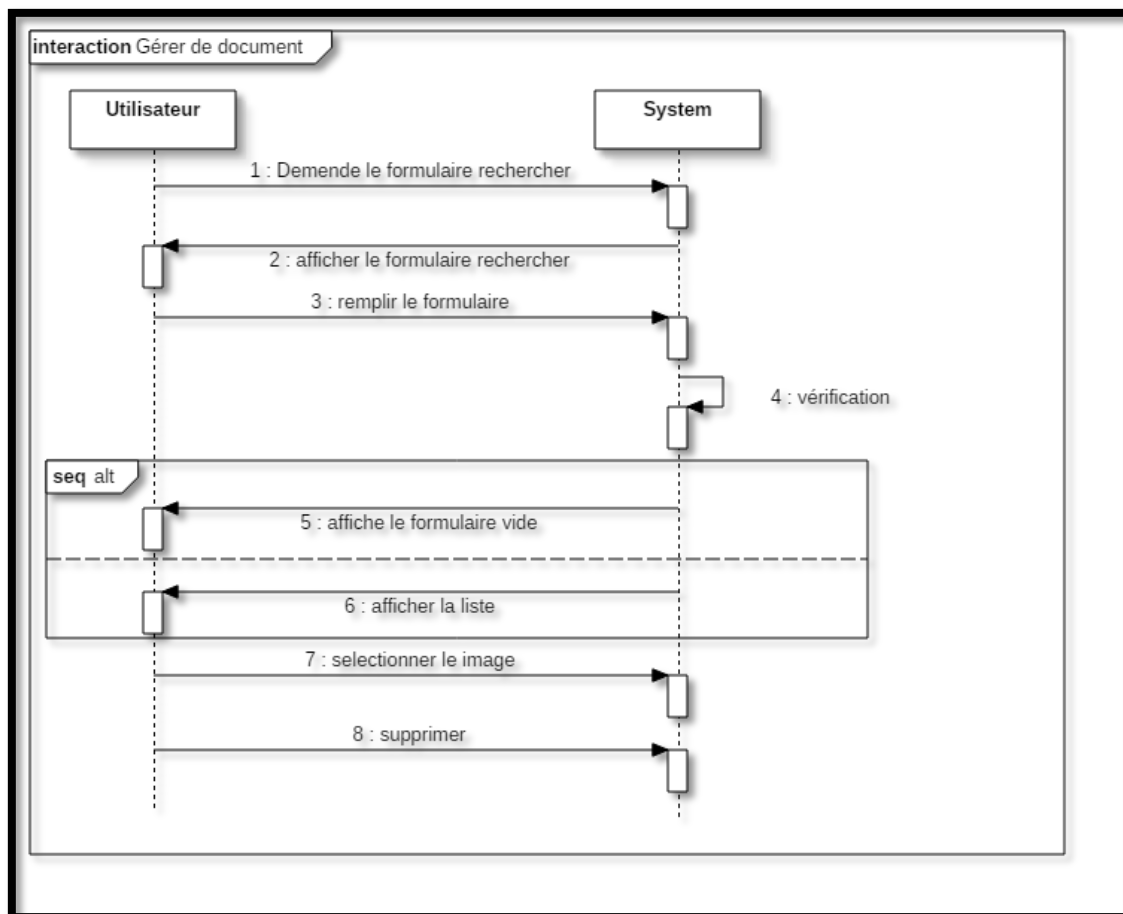


Figure .II.18: Diagramme de séquence " supprimer".

6. Diagramme de classe

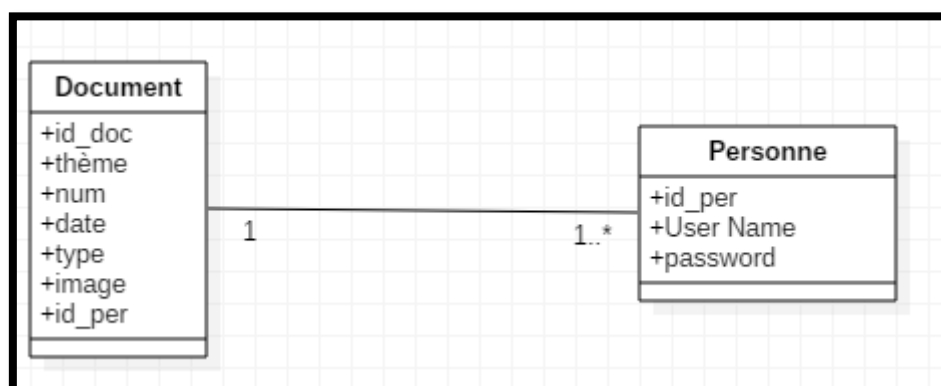


Figure II.19 : Diagramme de classe.

6.1. Description du modèle des classes

Classes	Attributs	Désignation	Type [Tille]
Personne	Id_per : int User : Varchar Password : Varchar	L'id personne Nom d'utilisateur Mot de passe	Int Varchar(50) Varchar(50)
Document	Id_doc : int Thème: Varchar Num :int Date :date Type :Varchar Image : image	L'id document Thème de document Numéro de document Date de document Type de document image	Int Varchar(50) Int(20) D Varchar(50) JPG

Tableau II.02 : Liste des classes.

7. Conclusion

Au cours de ce chapitre, nous avons entreprend la conception de notre système, et ainsi sa modélisation en termes des diagrammes de cas d'utilisation, séquence et de classe. Nous allons voir dans le chapitre suivant l'implémentation de notre système en détaillant tous les outils nécessaires pour cette réalisation.

Chapitre III

Développement du système

1. Introduction

Après avoir la conception de notre système dans le chapitre précédent et les différents diagrammes construits, dans ce chapitre on va détailler l'implémentation de notre application, on va voir également l'environnement de travail utilisé pour développer ce projet, puis nous présenterons les interfaces principales de notre logiciels.

2. Environnement de travail

2.1. Langage de programmation "C#"

Visual studio 2017 est un environnement de programmation riche en outils comportant toutes les fonctionnalités nécessaires pour créer des projets C# de toute taille

Le langage star de la nouvelle version de Visual studio et de l'architecteur .Net est C#, un langage dérivé du C++. il reprend certaines caractéristique des langage apparus ces dernière années et en particulier de Java .C# peut être utilisé pour créer, avec une facilité incomparable, des application Windows et Web. C# devient le langage de prédilection d'ASP.NET qui permet de créer des page Web dynamiques avec programmation coté serveur.[13]

❖ Les points forts de C#

- libération automatique des objets.
- le compilateur C#, comme Java, produit un p-code (bytecode) MSIL, Microsoft Intermediate Language, destiné à fonctionner dans une machine virtuelle (managed execution environment).
- C# a une syntaxe plus familière pour les programmes C, C++, Java et perl.
- la syntaxe élégante et économique est empruntée à C et C++.
- Grandes capacités d'introspection de classe (réflexion).

2.2..NET Framework

Les programmes en C# s'exécutent sur le .NET Framework, composant intégral de Windows qui inclut un système d'exécution virtuel appelé : "Common Language Runtime" (CLR) et un jeu unifié de bibliothèque de classe. Le CLR est l'implémentation commerciale de "l'infrastructure du langage commun" (CLI) de Microsoft, norme internationale constituant la base de toute création d'environnement d'exécution et de développement et assurant le fonctionnement homogène des langage et des bibliothèque.[14]

2.3. Microsoft Visual studio .Net

Visual studio .NET est un jeu complet d'outils de développement permettant de générer des applications Web ASP ,des service Web XML , des applications bureautique et des applications mobiles. Visual Basic .NET, Visual C# .NET , Visual C++ .NET et Visual J# .NET utilisent tous le même environnement de développement intégré(IDE,integrated développement environnement),qui leur permet de partager des outils et facilite la création de solutions faisant appel à plusieurs langages. Par ailleurs, ces langages permettent de mieux tirer parti des fonctionnalités du .NET Framework, qui fournit un accès à des technologies clés simplifiant le développement d'application Web ASP et de services Web XML.[15]

2.4.SGBD:(système de gestion de base de donnée)

SQL Server 2008 est un système de gestion de bases de données relationnelles ,produit par Microsoft . le stockage , la manipulation et l'analyse de ces données se font au sein de son moteur de bases de données . Ce service permet la réalisation de nombreuses applications , requêtes ,et transactions ,notamment grâce au langage T-SQL(Transact-SQL) .[16]

❖ La mise en œuvre de instances et des bases de données SQL Server

Une instance créée par SQL Server est un ensemble de binaires isolé les services du moteur de base de données SQL Server à partir d'autre applications. Fichiers de base de données distincts sont également utilisés .Les binaires fournissent les fichiers exécutables, fichiers de réseau et les outils administratifs nécessaires à l'exécution, une base de données sur un système Windows . Lorsqu'une instance est exécuté, il rassemble ses valeurs de configuration et les pools de mémoire alloue (tampons) dans la RAM et génère des discussions de fournir le CPU avec jeux d'instructions.

Un seul ordinateur peut avoir plusieurs instances SQL Server . L'objectif principal derrière l'architecture multi_instance est de contribuer à l'isolement de ressources pour les différents types d'activité de base de données au sein du même système physique . Chaque instance utilise indépendamment de la mémoire et de ressources processeur disponibles sur le système. SQL Server 2008 prend en charge jusqu'à 50 instances sur les édition Standard et Entreprise.[17]

3. Présentation d'application

Dans cette partie, nous présentons les différentes interfaces graphiques, sous forme des imprimés écran, de notre Logiciel.

3.1. Interface de « démarrage »

Cette interface représente une aide partielle pour accélérer le programme en chargeant toutes les interfaces.



Figure .III.1 :L' interface principale de l'application.

3.2. Interface « d'authentification »

Pour des raisons de sécuriser les différentes fonctionnalités et informations, cette interface permet d'authentifier les différents utilisateurs.

- ✓ Si le mot de passe correcte, il connecte au le programme.
- ☒ Sinon le système signale un erreur.

La figure.III.2 ci-dessous représente l'interface de l'authentification.

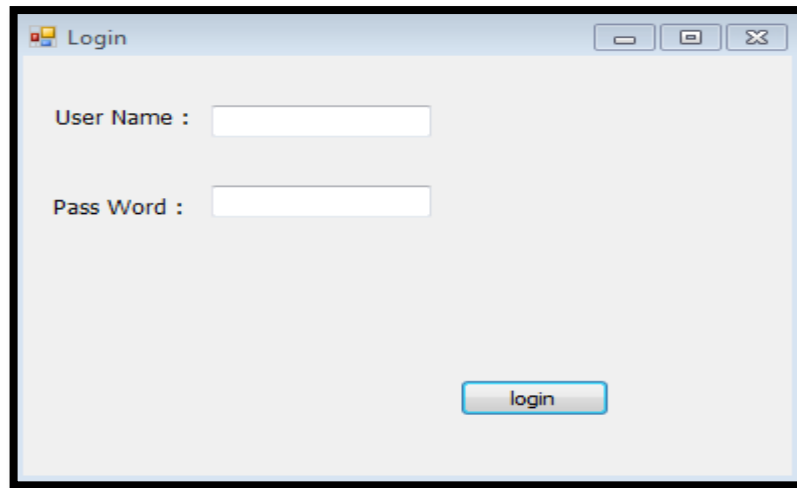


Figure.III.2: L'interface de l'authentification de l'application.

3.3.Interface «scanner »

Pour scanner le document , l'utilisateur clique sur "login" dans l'interface authentification, puis se remplis leur informations

✓ Si les information correctes ,le système sauvegarde les information

☒ Sinon il informe l'utilisateur sur les erreurs

La figure .III. 3: présente l'interface de scanner.

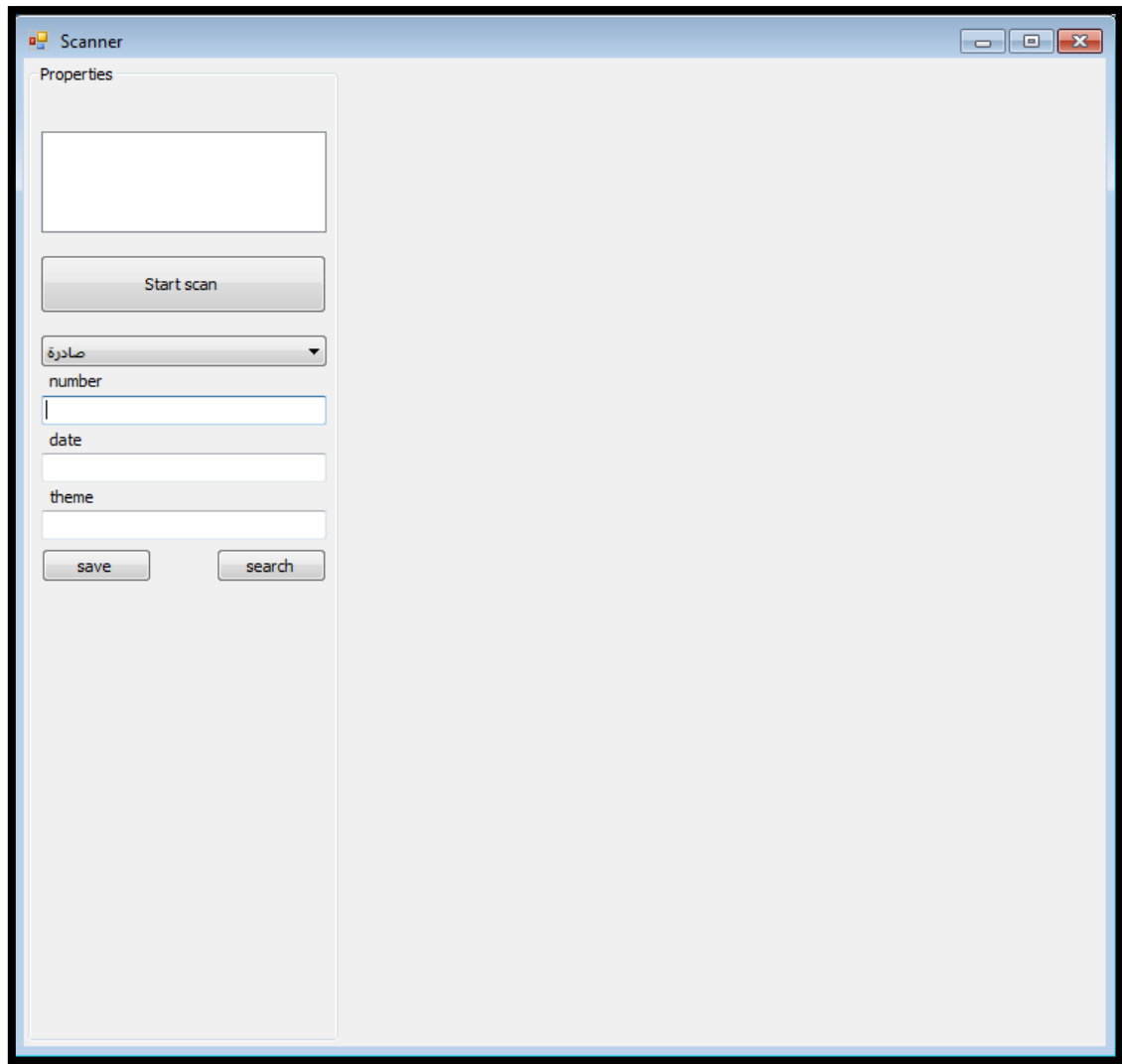


Figure .III. 3:Interface de scanner

Après le scanner de document, il affiche l'image, nombre et la date

La figure .III. 4: présente l'interface après scanner le document.

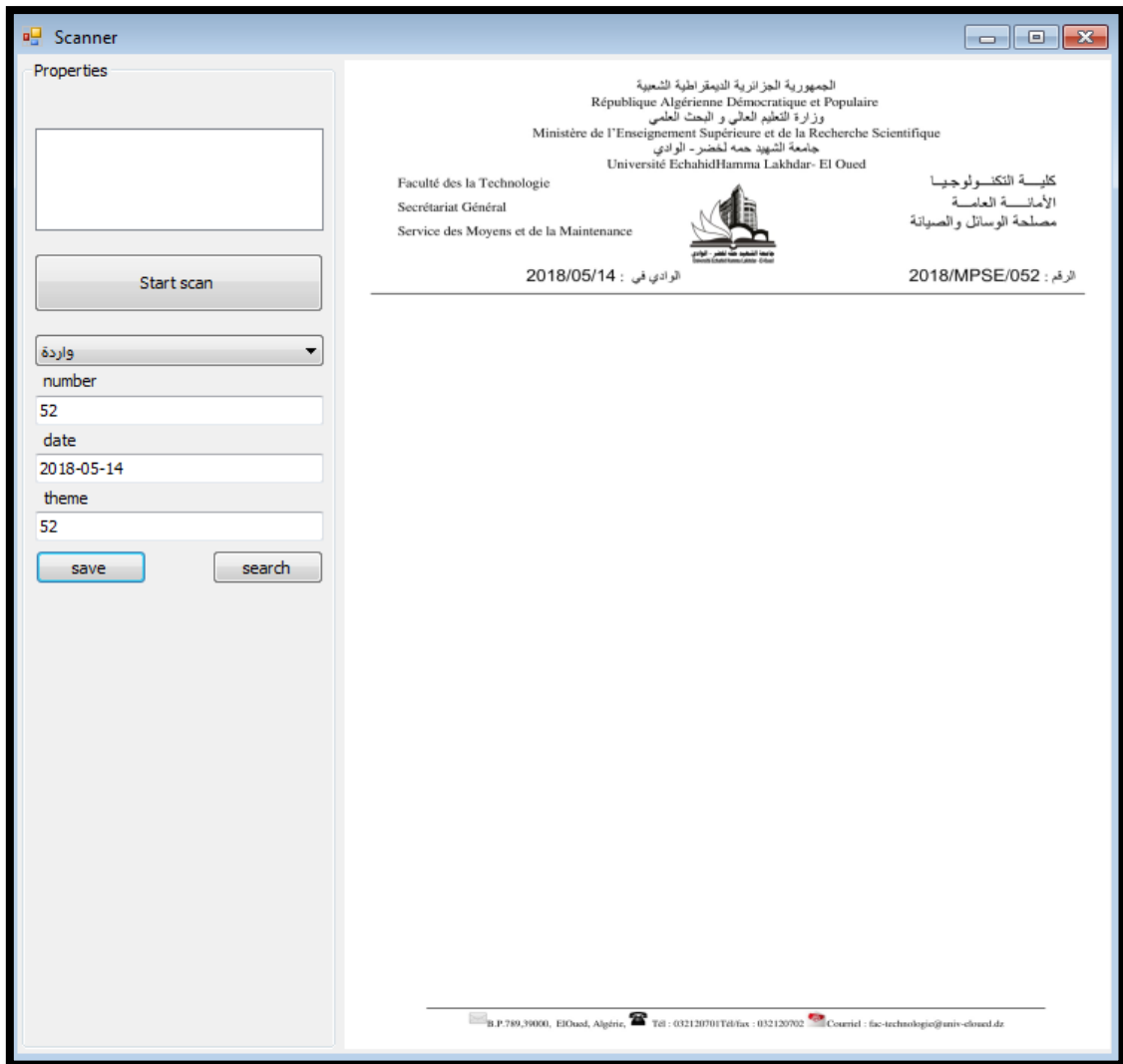


Figure.III.4: L'interface Après scanner le document

3.4. Interface « recherche »

Pour rechercher à un document , l'utilisateur clique sur "login" dans l'interface authentification, puis se remplit leurs informations .

- ✓ Si les informations sont correctes ,le système sauvegarde les informations
- ☒ Sinon il informe l'utilisateur sur les erreurs

Et l'utilisateur clique sur "search" dans l'interface scanner ,il remplit la date ou le nombre et affiche la liste de documents

La figure .III. 5: présente l'interface de recherche de documents.

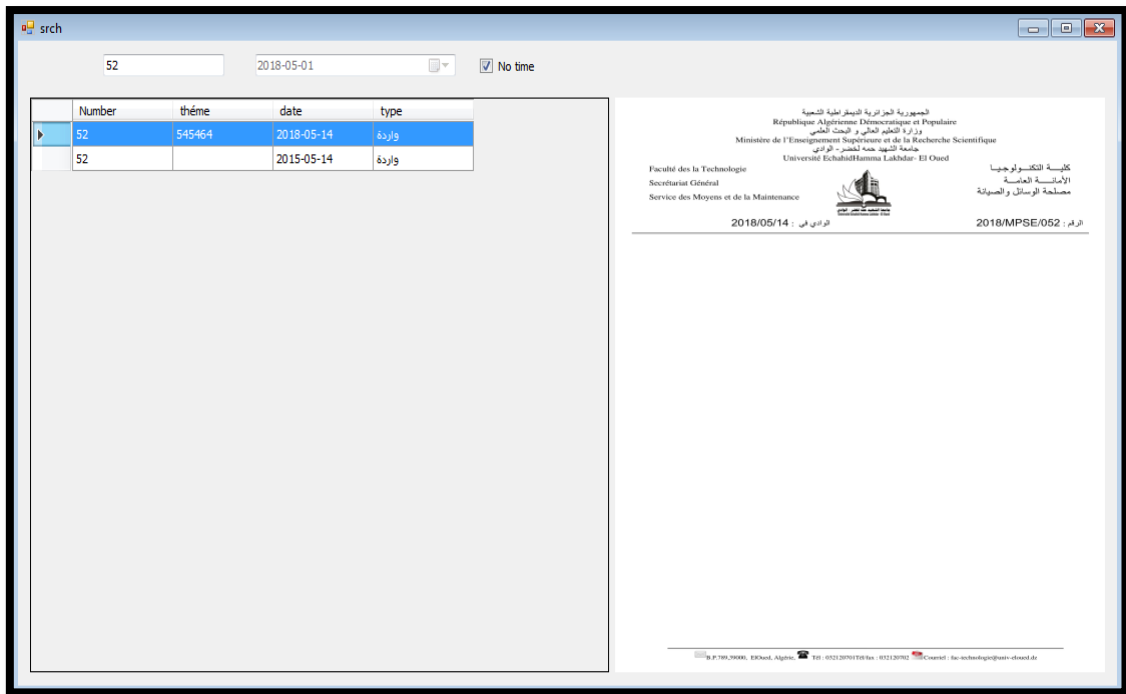


Figure .III. 5: Interface recherche .

3.5. Interface « imprimer ou supprimer»

Après la recherche l'utilisateur clique sur "login" dans l'interface authentification, puis se remplis leur informations .

- ✓ Si les information correctes ,le système sauvegarde lei information
- ☒ Sinon il informe l'utilisateur sur les erreurs

Et l'utilisateur clique sur "nombre ou date" dans l' interface "search" après

Clique sur " supprimer " ou "imprimer"

La figure .III. 6: présente l'interface supprimer ou imprimer le document.

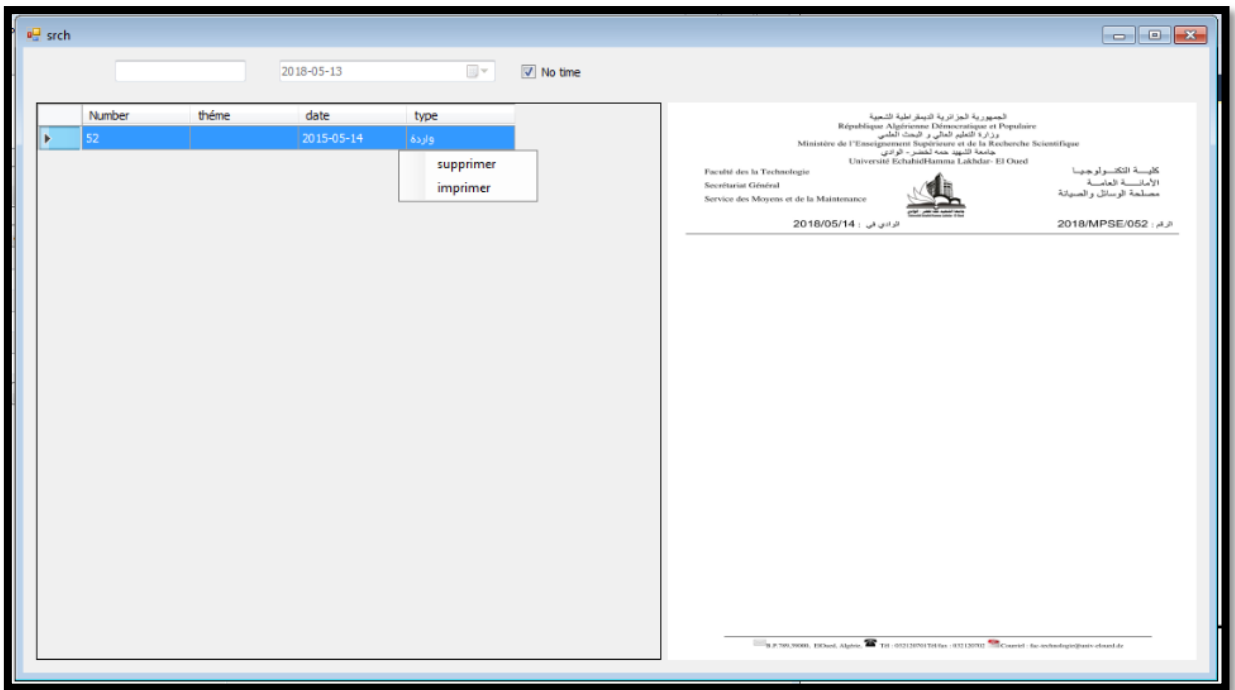


Figure.III.6: L'interface imprimer ou supprimer le document

4. Conclusion

Dans ce chapitre, nous avons présenté la partie de réalisation de notre projet, et nous avons illustré les interfaces les plus importantes de notre application.

Conclusion générale

L'objectif principal de notre travail est le développement d'une application de Classification de documents numérisés via scanner en utilisant la technique de reconnaissance optique de caractères (OCR).

Notre mémoire est composé de deux parties: une partie théorique qui englobe deux chapitres et une partie pratique consacré dans le dernier chapitre.

Dans le premier chapitre nous avons donnée une vue générale sur le cadre général de notre projet, puis nous avons étudié quelques solutions proposées dans ce contexte afin d'extraire les points forts et faibles. Dans le deuxième chapitre, nous avons abordé la conception de notre solution. dans le troisième chapitre nous avons traduit notre modélisation conceptuelle en une implémentation tout en montrant les différentes interfaces obtenues

En fin de notre étude, nous espérons avoir répondu aux besoins et soucis des utilisateurs.

Référence

- [1]: VISION Wikipédia. Vision par ordinateur [en ligne].
<https://fr.wikipedia.org/wiki/Vision_par_ordinateur> (consulté le 13/03/2018)
- [2]:DOMAIN Numérisations [en ligne]
.<<http://numerisations.blogspot.com/2012/05/blog-post.html>> (consulté le 13/03/2018)
- [3]:GESTION https://fr.wikipedia.org/wiki/Gestion_électronique_des_documents
- [4]:SCAN<https://ar.wikipedia.org/wiki>
- [5]: OCR Wikipédia Reconnaissance optique de caractères [en ligne]. <
https://fr.wikipedia.org/wiki/Reconnaissance_optique_de_caractères> (consulté le 19/04/2018)
- [6]:TESSERACT <https://github.com/tesseract-ocr/tesseract/wiki>
- [7]:FAC et Départ <http://www.univ-eloued.dz/>
- [8]:ARCHIVAGE
<https://abarifi.wordpress.com/2012/04/20/%D8%A7%D9%84%D8%A7%D8%B1%D8%B4%D9%8A%D9%81/>
- [9]:2TUP Processus de développement en Y(Processus 2TUP) [en ligne]
<<https://ar.scribd.com/doc/49697489/Processus-de-Developpement-Y-Processus-2TUP>>
(consulté le 16/03/2018)
- [10]:UML... Laurent AUDIBERT,'UML2'Institut Universitaire de Technologie de Villetaneuse-Département Informatique ,édition 2007-2008
- [11]:AVANTAGE ET CONVI Conception et réalisation d'une application d'orientation des étudiant des 1ST , Université d'El oued –Département des mathématique et Informatique License informatique , édition 2011-2012
- [12]:Modélisation des besoin...: F .Juliard UML Unified Méthode Language ,Journal Université de Bretagne Sud UFR SSI-IUPV,2001-2001.

[13]:LANGAGE G.LebLANC/Livre C# et .NET version 2/Edition EYROLLES /2006

[14]:NET FRAM<http://msdn.microsoft.com/fr-fr/library/zlzx9t92%28VS.80%29.aspx>.03/05/2018

[15]:VISUEL <http://msdn.microsoft.com/fr-fr/library/aa291755%28v=VS.71%29.aspx>.03/05/2018.

[16]:SGBD VALENTIN Pauline/Introduction à la B.I. avec SQL Server 2014/

[17]: MISE EN OUEVRE Microsoft /SQL server 2014/ SQL Server Technical Article