



HAL
open science

Datation automatique de photographies à partir de caractéristiques textuelles et visuelles

Paul Martin

► **To cite this version:**

Paul Martin. Datation automatique de photographies à partir de caractéristiques textuelles et visuelles. Informatique [cs]. Université Caen Normandie, 2015. Français. NNT: . tel-01260365

HAL Id: tel-01260365

<https://hal.science/tel-01260365v1>

Submitted on 22 Jan 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Université de Caen Normandie

École doctorale SIMEM

Thèse de doctorat

présentée et soutenue le : 17/12/2015

par

Paul MARTIN

pour obtenir le

Doctorat de l'Université de Caen Normandie

Spécialité : Informatique et applications

**Datation automatique de photographies
à partir
de caractéristiques textuelles et visuelles**

Directeur de thèse : Marc SPANIOL

Codirecteur de thèse : Antoine DOUCET

Jury

M. Josep LLADÓS	Full Professor	Univ. Autònoma de Barcelona	<i>Rapporteur</i>
M. Philippe RIGAUX	Professeur des Universités	CNAM	<i>Rapporteur</i>
M. Patrice BELLOT	Professeur des Universités	Aix-Marseille Université	<i>Examinateur</i>
M. Frédéric JURIE	Professeur des Universités	Univ. Caen Normandie	<i>Examinateur</i>
M. Pascal BULÉON	Professeur des Universités	Univ. Caen Normandie	<i>Membre invité</i>
M. Marc SPANIOL	Professeur des Universités	Univ. Caen Normandie	<i>Directeur de thèse</i>
M. Antoine DOUCET	Professeur des Universités	Univ. La Rochelle	<i>Codirecteur de thèse</i>

Résumé

Dans cette thèse, la problématique que nous traitons est celle de la datation automatique de photographies anciennes à partir d'attributs issus de l'image et du texte qui l'entoure.

Au cours de ces derniers siècles, la photographie et les métadonnées ont subi de nombreuses mutations. Dans une vocation informative et contextualisante, nous reportons nos observations au regard de ces avancées technologiques qui permettent de comprendre l'évolution des méthodes de datation de photographies, historiquement manuelles et récemment devenues automatiques.

Nous nous sommes intéressés à une notion fondamentale et intrinsèque de la datation : le caractère ordinal du temps. Nous avons donc étudié les méthodes de classification ordinale et les mesures d'évaluation associées.

Pour cette thèse, nous avons construit une large collection de pages web associant clichés anciens et contenus textuels. Celle-ci a pour vocation de permettre l'entraînement et l'évaluation des modèles proposés. Ceci représente une avancée dans ce domaine car aucun jeu de données comparable n'existait auparavant.

Nous proposons un modèle original que nous avons développé et évalué sur des bases de photographies pour les dater grâce à leur contenu visuel. La particularité majeure de notre approche est l'exploitation pour la première fois du caractère ordinal du temps.

Enfin, nous explorons la datation de photographies par l'usage des données textuelles environnantes. Notre démarche imite le comportement humain utilisé lors de la datation de photographies en contexte. Ainsi nous proposons pour la première fois une méthode automatique permettant de dater des clichés par l'usage de l'environnement textuel.

Abstract

In this thesis, we address the problem of automatic dating of old photographs using attributes from image and surrounding text.

Over the last centuries, photography and metadata have undergone many mutations. In an informative and contextualizing vocation, we report our observations regarding these technological advances. This information helps to understand the evolution of photographs dating methods that have historically been studied manually before becoming recently automatic.

We focused in a fundamental concept of dating : the ordinal nature of time. To this extent we studied ordinal classification methods and associated evaluation measures.

In this thesis, we have built a large collection of web pages combining old photographs and text content. This collection aims at allowing both training and evaluation of the proposed models. This represents an advance in the field of photographs dating because no comparable data set previously existed.

We propose an original model that we have developed and evaluated on image databases to date the photographs with their visual content. The major feature of our approach is the exploitation of the ordinal structure of time. This is to our knowledge the first use of this intrinsic characteristic of photographs dating.

Moreover, we explore the dating of photographs by the use of the surrounding text data. In our approach, we mimic the human behavior used in the context of photographs dating. Thus we propose for the first time an automatic method for the dating of photographs by the use of the textual environment.

Je tiens à exprimer mes sincères remerciements à mes encadrants, Messieurs les Professeurs d'Université Antoine DOUCET, Frédéric JURIE et Marc SPANIOL, qui ont su m'orienter dans les bonnes directions et me conseiller tout au long de cette thèse. Sans leur aide précieuse et leurs conseils avisés, les travaux que nous présentons dans ce manuscrit n'auraient pu être ainsi menés. Leurs apports, tant sur le plan scientifique que technique ont été essentiels à l'élaboration de cette thèse et ce, y compris dans leurs relectures.

Merci.

Ma gratitude va également à Messieurs Josep LLADÓS, Philippe RIGAUX, Patrice BELLOT et Pascal BULÉON qui ont accepté d'être membres de mon jury de thèse.

De plus, j'aimerais remercier toutes les équipes IMAGE et HULTECH¹ avec qui j'ai eu la chance de pouvoir échanger et avancer sur ce sujet, notamment Messieurs Gaël DIAS, Ashish GUPTA, Alexis MIGNON, José MORENO et Gaurav SHARMA. J'aimerais tout particulièrement adresser mes remerciements à Monsieur Gaël LEJEUNE pour son implication et son dévouement dans la relecture approfondie de mon manuscrit.

Je tenais à remercier aussi le GREYC², l'école doctorale SIMEM³ ainsi que mes collègues de travail des départements administratifs et du support technique, notamment Mesdames Virginie DESNOS-CARREAU, Arielle PERRETTE, Véronique ROBERT, Véronique WATROBA et Messieurs Pierre BLONDEAU, Davy GIGAN et Renaud RABEIL.

Sans oublier le soutien financier de la région Basse-Normandie grâce auquel, les travaux au cours de ces 3 années de thèse ont pu se développer.

Mes remerciements vont enfin à ma famille, à mon amie, à mes amis et à toutes celles et ceux qui m'ont apporté leur aide durant toutes les étapes de cette expérience.

Merci à vous.

– Paul MARTIN.

1. HUmAn Language TECHnology

2. Groupe de Recherche en Informatique, Image, Automatique et Instrumentation de Caen

3. Structures, Information, Matières et Matériaux

Table des matières

Résumé	I
Abstract	III
1 Introduction	5
1.1 Vers une société de l'information et de l'image	9
1.2 Datation d'images et applications	10
1.2.1 Hypothèses du lien entre le texte et l'image	12
1.2.2 Dimensions de l'information visuelle	13
1.2.3 Recherche d'information temporelle	15
1.2.4 Exemples d'usage de l'information temporelle	16
1.3 Contributions et plan du manuscrit	18
2 Historique de la photographie et des métadonnées	21
2.1 Photographie	21
2.1.1 Photographie traditionnelle	22
2.1.2 Photographie digitale	23
2.2 Métadonnées	25
2.2.1 Métadonnées traditionnelles	25
2.2.2 Métadonnées digitales	27
2.3 Observations	30
3 Datation de photographies : des méthodes manuelles aux automatiques	31
3.1 Origines de la datation de photographies : méthodes manuelles	31
3.1.1 Photographie à travers le temps	32
3.1.2 Démarche de l'expertise manuelle pour la datation de photographies	33
3.2 Méthodes automatiques de datation des photographies numérisées	35
3.2.1 Approches par détection de filigranes	35
3.2.2 Approches utilisant des caractéristiques visuelles	38
3.3 Méthodes automatiques pour la datation des écrits	41
3.3.1 Approches visant à retrouver la date de création de l'œuvre	41
3.3.2 Approches visant à comprendre la période concernée par l'œuvre .	42
3.4 Propositions vis-à-vis des méthodes existantes	44
4 Classification ordinale	45
4.1 Fondements théoriques	46
4.1.1 Classification	47
4.1.2 Régression	47
4.1.3 Classement	48

4.2	Définitions	49
4.2.1	Approche générale	49
4.2.2	Familles d'approches	50
4.2.3	Approche par décomposition en sous-problèmes binaires	51
4.3	Méthodes d'évaluations de la classification ordinale	56
4.3.1	Mesures usuelles en classification	57
4.3.2	Variante tolérant certaines erreurs	58
4.3.3	Mesures issues de la régression	58
4.3.4	Améliorations des mesures fondées sur la distance	60
4.3.5	Mesures utilisant l'ordre relatif	61
4.3.6	Mesure hybride	63
4.3.7	Utilisation des méthodes d'évaluation en apprentissage supervisé	64
4.4	Jeux de données ordinaux	65
4.4.1	Jeux de données issus de la régression	66
4.4.2	Jeux de données de synthèse générés	67
4.4.3	Jeux de données naturellement ordinaux	67
4.5	Discussion	69
5	Constitution d'un jeu de données de photographies datées et contextualisées	71
5.1	Recherche de bases de photographies datées et contextualisées	72
5.1.1	Bases d'images datées	73
5.1.2	Expérimentations à la réalisation d'un premier jeu de données	77
5.2	Constitution du corpus web-photo 1900-2000	81
5.2.1	Collecte d'informations issues du web	82
5.2.2	Annotation du corpus	83
5.2.3	Élaboration d'un sous-corpus de qualité : web-photo 1900-2000	84
5.3	Conclusion	86
6	Datation visuelle	89
6.1	Introduction	90
6.2	Nouvelle approche de datation visuelle	91
6.3	Vers une nouvelle méthode de classification ordinale : <i>Avant contre Après</i>	93
6.3.1	Ré-étiquetage des sous-classifieurs binaires	94
6.3.2	Apprentissage des sous-classifieurs binaires	94
6.3.3	Prédiction à partir des sous-classifieurs binaires	96
6.4	Expérimentations	99
6.5	Conclusion	102
7	Datation textuelle	107
7.1	Introduction	108
7.2	Démarche conceptuelle	110
7.3	Chaîne de traitement de la réconciliation temporelle	114
7.3.1	Approche stricte	116
7.3.2	Approche souple	117
7.4	Expériences et résultats	120
7.5	Conclusion	123
	Conclusions et perspectives	125

Liste des illustrations	137
Liste des tableaux	140
Références bibliographiques	155
A Frises chronologiques	157
A.1 Frise chronologique représentant l'évolution du temps en géologie	157
A.2 Frises chronologiques retraçant l'évolution de la photographie	158
B Impacts du numérique sur l'économie de la photographie	163
B.1 Déclin des ventes de pellicules et d'appareils	163
B.2 Déclin des développements de pellicules	164
B.3 Faillite de <i>Kodak</i> en 2012	164
C Spécifications de la norme XMP	167
C.1 Category : Schema properties are internal or external	167
C.2 EXIF Dates	168
C.3 Dublin Core Schema Date	168
C.4 XMP Basic Schema	168
C.5 Photoshop Schema	168
C.6 XMP Schemas (EXIF Schemas)	169
C.6.1 tiff prefix schema	169
C.6.2 exif prefix schema	169
D Véracité des données EXIF et XMP extraites à partir de 150 000 images	171
D.1 Statistiques données EXIF et XMP	172
E Autres jeux de données ordinaux	177
E.1 Jeux de données naturellement ordinaux	177
F Captures d'écrans et présentations des bases d'images découvertes	181
G Informations complémentaires de notre jeu de données issu de flickr	189
G.1 Présentation d'un document de notre base	189
G.2 Présentation de notre système d'annotation d'images	191
G.3 Statistique de répartition des photographies	192
H Informations complémentaires de notre jeu de données issu du web	195
H.1 Liste des 50 villes pour lesquelles nous avons effectué des recherches	195
H.2 Présentation d'un document de notre base	197
H.3 Répartition des formats de photographies à travers les décennies	201
I Observations sur le protocole expérimental	207
J Comparaison de méthodes de classification ordinale	209

Chapitre 1

Introduction

Sommaire

1.1 Vers une société de l'information et de l'image	9
1.2 Datation d'images et applications	10
1.3 Contributions et plan du manuscrit	18

Au cours de cette thèse, nous nous sommes intéressés à la problématique de la prédiction automatique de la date de prise de vue d'une photographie ancienne à partir de son contenu visuel, du texte qui l'entoure ou de toute autre information disponible pour cette photographie comme par exemple les métadonnées.

Cette problématique est présentée dans l'illustration 1.1 où nous montrons une photographie ancienne insérée dans une page web et pour laquelle la date, ici l'année 1959, peut être déduite au moyen du texte qui la précède. Le contenu de l'image, comme par exemple les modèles de voitures présents, permet également d'inférer des informations sur la période sans toutefois être aussi précis qu'avec l'information textuelle.

Cette thèse sera donc liée à 2 disciplines scientifiques : la vision par ordinateur pour ce qui est de l'interprétation du contenu des images et le traitement automatique des langues pour ce qui est de l'interprétation des informations textuelles.

Le traitement automatique des langues et la vision par ordinateur s'inscrivent dans une problématique plus générale qui est celle du traitement de l'information et des contenus multimédias. Il s'agit de disciplines importantes à notre époque, où l'acquisition d'information numérique devient prédominante, permanente et grandissante : l'homme veut conserver et stocker toujours plus de données (textes, images, vidéos, *etc.*). Si le stockage de l'information pose peu de difficultés particulières grâce aux progrès technologiques, l'exploitation de celle-ci reste un problème de taille et présente à l'heure actuelle de réels enjeux pour l'avenir. C'est dans ce contexte que se situe cette thèse.

HACE 53 AÑOS

El Rally Barcelona Sitges cumple 53 años

En febrero de 1959, Samaranch despedía en la plaza de Sant Jaume a los primeros 23 automóviles. Todos compartían objetivo: ¡resistir! Todos habían sido fabricados antes de 1920. ¿Cuántos saldrán este año? Atentos a la... ¡Salida!

Hemeroteca | 08/02/2012 - 00:01h



DOCUMENTACION

ILLUSTRATION 1.1: Exemple d'image ancienne insérée dans une page web : l'objectif est de prédire automatiquement la date de prise de vue de cette image en se basant à la fois sur le contenu visuel de l'image et sur le texte qui l'entourne. (source lavanguardia.com)

Dans le contexte général du traitement de l'information, nos travaux s'inscrivent dans le domaine plus spécifique de la compréhension automatique de la temporalité dans les images. Ceci représente une problématique complexe, aujourd'hui encore largement ouverte. Nous nous appuyons, comme mentionné ci-dessus, sur deux dimensions : l'analyse du contenu des images et l'analyse du texte qui l'entoure, qui toutes deux soulèvent de nombreuses questions, en particulier liées à la difficulté de séparer l'aspect et le contenu. C'est ce que nous allons développer dans les paragraphes qui suivent.

En ce qui concerne l'image, les informations, liées à l'aspect mais aussi au contenu, influencent la perception temporelle des photographies. L'humain est cependant capable de les distinguer et de percevoir la concordance ou la discordance entre ces informations. Nous présentons ce point dans l'illustration 1.2.

En revanche, pour une machine, ces deux notions sont ambiguës et leur différenciation sera donc un enjeu majeur pour la datation.

Il existe en pratique une multitude de combinaisons possibles entre les aspects photographiques et les contenus temporels photographiés. Certaines photographies paraissent difficiles à dater, comme par exemple les clichés avec un contenu contemporain mais un aspect ancien ; le vieillissement artificiel de photographies étant un thème « à la mode » (voir par exemple l'application *instagram*¹, qui propose ce type de transformation). En ce qui concerne les photographies présentes sur la toile, elles correspondent majoritairement à des clichés réalisés avec des appareils contemporains de leurs sujets.

	Contenu ancien	Contenu contemporain
Aspect ancien		
Aspect contemporain		

ILLUSTRATION 1.2: Opposition entre contenu et aspect d'une photographie selon les époques.

Du point de vue de la vision par ordinateur, le contenu visuel est difficile à extraire, reconnaître et classifier. De plus, les méthodes modernes reposent sur des principes d'apprentissage statistique et requièrent des bases de données volumineuses² pour pouvoir être entraînées. À l'inverse, l'analyse de l'aspect de photographies est une notion plus générale que le contenu et sera peut être plus simple à appréhender. Pour cette raison, nos travaux se concentreront d'abord sur la datation automatique à partir de l'analyse de l'aspect d'une image (e.g. ton jaunâtre, faible résolution, brillance), avant de se pencher également sur l'extraction du contenu.

1. <https://instagram.com> dernier accès août 2015

2. Voir par exemple KRAPAC, ALLAN, VERBEEK et JURIE (2010).

Nous pouvons observer un parallèle intéressant entre la datation visuelle et la datation textuelle. L'aspect d'une photographie (la forme) peut être apparenté au style du texte (e.g. de la page web). En effet, le style linguistique peut être considéré comme une technique que l'on utilise pour exprimer l'information du contenu. Parallèlement, le vocabulaire textuel (le fond) s'apparente lui plutôt au contenu d'une photographie. Les mots émergent à mesure que les objets apparaissent dans le temps (e.g. le mot *iPhone* à partir de 2007). De la même façon que nous avons opposé le contenu et l'aspect pour les images, nous montrons dans le tableau 1.1, l'opposition entre le style et le vocabulaire textuel. Cette observation est selon nous, une source d'information importante pour la datation d'images à partir de contenus textuels.

	Vocabulaire ancien	Vocabulaire contemporain
Style ancien	« affublé d'une redingote »	« affublé d'un short baggy »
Style contemporain	« sapé avec une redingote »	« sapé avec short baggy »

TABLEAU 1.1: Opposition entre vocabulaire et style textuel selon les époques.

Du point de vue du traitement automatique des langues, la tâche de datation des textes a été particulièrement étudiée à travers des modèles temporels du langage³, qui reposent sur l'analyse de l'évolution du vocabulaire dans le temps.

De ces observations, il apparaît que pour l'image il est plus facile d'accéder à l'information liée à l'aspect alors qu'à l'inverse, avec le texte, il est plus facile d'accéder au vocabulaire (s'apparentant au contenu). Il semble que ces 2 sources soient complémentaires, et cette complémentarité nous renforce dans l'idée d'étudier conjointement ces deux domaines dans le cadre de la datation de photographies.

Les travaux que nous avons menés dans cette thèse se tournent résolument vers les photographies anciennes, « nées sous l'ère analogique » pour lesquelles une digitalisation a pu être effectuée. Pour bon nombre de ces photographies, la date n'est pas renseignée précisément et l'extraire automatiquement ouvre la porte à des applications de recherche ou d'indexation de documents. Notre objectif vise donc, plus particulièrement, à retrouver l'information qui n'est pas transférée lors de la numérisation. En effet, le passage de l'analogique au numérique ne transfère généralement pas les métadonnées relatives aux photographies originales (par exemple, la date inscrite au dos du cliché).

3. Voir par exemple JONG, RODE et HIEMSTRA (2005).

1.1 Vers une société de l'information et de l'image

L'arrivée du web 2.0 au milieu des années 2000 a favorisé l'interconnexion entre les utilisateurs d'Internet. En effet, les utilisateurs sont sans cesse plus nombreux à accéder à Internet, mais surtout plus connectés entre eux et interagissent en masse dans ce qui a été nommé « le web social » sans doute du fait de sa démocratisation et son omniprésence dans les sociétés. Nous sommes passés d'une société de l'information où l'on accédait au contenu dans le web statique dit « 1.0 » à une société de la communication dans laquelle les interactions et l'apport d'information tiennent une place prépondérante : les utilisateurs qui étaient seulement consommateurs sont devenus aussi acteurs de la toile. La conséquence de cette évolution est de produire une quantité sans cesse grandissante de contenus parmi lesquels nous pouvons citer les images, les textes, les vidéos, les commentaires, les métadonnées, les liens, *etc.* Le nombre de documents numériques contenant des images augmente donc de façon considérable.

Un certain nombre d'acteurs dominants se sont positionnés sur ce créneau du web social, offrant des services pour lesquels les photographies jouent une place prépondérante :

- *flickr*⁴ propose gratuitement l'hébergement de photographies à hauteur de 1 téra-octet (To), soit l'équivalent de 200 000 images de 5 Mo chacune ;
- *instagram*⁵ est un réseau social de partage d'images qui permet d'appliquer des filtres sur les images (aspect ancien simulé...) ;
- *facebook*⁶ est un réseau social servant de journal électronique (photographies de famille, événements sociaux, *etc.*) ;
- *snapchat*⁷ est un réseau social dans lequel il est possible de partager une photographie numérique auto-destructible après un laps de temps défini (la photographie digitale comme bien de consommation) ;
- *twitter*⁸ est un réseau social basé sur l'envoi de mini messages de tailles inférieures à 160 caractères, auxquels il est aussi possible d'ajouter des images.

Certains musées à travers le monde ont également procédé à la numérisation de tout ou partie de leurs collections et se positionnent eux-aussi comme des acteurs majeurs de ce monde numérique. Leur objectif est à la fois de rendre accessible aux utilisateurs d'internet le catalogue des œuvres en leur possession mais aussi d'enrichir leurs connaissances grâce à la myriadisation⁹. Les informations obtenues de cette manière présente un certain intérêt mais leur validité reste parfois à vérifier

4. <http://www.flickr.com> dernier accès août 2015

5. <http://www.instagram.com> dernier accès août 2015

6. <http://www.facebook.com> dernier accès août 2015

7. <http://www.snapchat.com> dernier accès août 2015

8. <http://www.twitter.com> dernier accès août 2015

9. Myriadisation (SAGOT, FORT, ADDA, MARIANI et LANG, 2011), en anglais : *Crowdsourcing*

La collection flickr « *The Commons* » illustre bien cette question de l'arrivée des musées dans le web 2.0. Lancé par flickr et la bibliothèque du congrès des USA le 16 janvier 2008, ce service implique actuellement plus de 100 institutions participantes dont l'Université de Caen Normandie ¹⁰. L'objectif est de partager les trésors cachés des photographies d'archives publiques du monde entier, de faciliter la recherche de photographies en laissant les utilisateurs ajouter des tags, des commentaires, et de permettre aux utilisateurs de contribuer au développement des contenus par la validation majoritaire de l'information fournie à force de recoupements ou de discussions au sein du système. Cette dynamique prouve que l'intérêt actuel se trouve dans les métadonnées car sans celles-ci cette masse d'informations a beaucoup moins de valeur en terme d'exploitation et de visibilité.

Ces nouveaux espaces de communication et de partage entre les utilisateurs dictent ainsi de nouveaux usages des technologies dont la photographie fait partie. D'une manière générale les utilisateurs se sont de plus en plus ouverts à partager des informations qui, il y a quelques années, étaient considérées comme privées. On citera en exemple le partage d'albums photographiques via les réseaux sociaux. Cette tendance n'est pas isolée car nous observons aussi l'apparition de phénomènes de « société digitale » tels que le « selfie ¹¹ » ou encore, la « *foodographie* ¹² ». Il est à noter que ces tendances sont tellement populaires qu'elles ont inspiré des recherches, notamment *Google Recherche* a travaillé sur le projet *im2Calories* ¹³ qui vise à calculer le nombre de calories contenues dans un plat photographié.

Les besoins, quant à eux, précèdent les avancées technologiques et prennent une place majeure dans l'évolution du web. Les utilisateurs se servent abondamment des moteurs de recherches, des encyclopédies en ligne et font de plus en plus confiance aux résultats proposés. Ainsi nous en sommes actuellement à un basculement du « web social » vers un « web sémantique ». Nous passons d'une société de la communication à une société de la connaissance tirant partie des métadonnées et de la structuration de l'information. Le « web sémantique » se veut hiérarchisé et largement interconnecté, car il enrichit la conscience et la connaissance collectives. Ainsi les données brutes subissent peu à peu des mutations rendant l'information intelligente et adaptée aux besoins populaires.

1.2 Datation d'images et applications

Nous définissons la datation comme l'action de situer le plus précisément possible dans le temps, par une date ou un âge : un fait, un objet ou une personne.

10. https://www.flickr.com/photos/universite_caen/albums dernier accès août 2015

11. *Selfie* : le fait de se photographier « soi-même » pour réaliser son autoportrait.

12. *foodographie* : le fait de partager un instantané de nourriture (e.g. un plat cuisiné, un cocktail, etc.) avant consommation.

13. Annoncé en Mai 2015 par Kevin Murphy (employé chez *Google*) au cours du sommet *Re-Work* sur le *Deep Learning*.

La précision de l'information définit ici un « pas » temporel, *i.e.* la plus petite unité temporelle indivisible dans laquelle la temporalité est exprimée. Ce pas va déterminer la granularité de l'information. Pour une machine, le temps est une réalité implacable définie clairement, strictement et régulièrement espacée. Pour l'homme le temps est une notion qui découle de sa perception, et de sa sensibilité. Cette question est et restera éperdument délicate car elle dépend avant tout de l'utilité que nous portons à l'information temporelle. Chaque domaine dispose donc d'un niveau de granularité adapté à son contexte. Les séparations temporelles définissant la granularité, ne sont en effet pas strictement régulièrement espacées dans le temps¹⁴ car notre représentation personnelle, émotionnelle et historique du temps reste subjective, et ne correspond pas à la réalité mais à l'usage que nous avons de l'information temporelle.

Par exemple, en géologie l'unité de mesure de la temporalité n'est pas la même qu'en course à pied. La proximité de l'événement influence aussi la précision que nous lui attribuons. Plus l'événement est proche temporellement plus nous allons lui attribuer une granularité détaillée. Par exemple l'événement : « le premier homme a avoir posé le pied sur la Lune » est dans les mémoires, daté de l'été de l'année 1969 mais il est en réalité daté : « 21 juillet 1969 UTC à 3h56 ». Si nous prenons un événement plus lointain temporellement : « Le couronnement de Charlemagne » est daté en l'an « 800 », ce qui lui confère une granularité plus grossière. En revanche pour des historiens, il est daté « le 25 décembre 800 ».

La temporalité et les granularités que l'on associe à une échelle temporelle peuvent se traduire par une structure hiérarchique entre les classes temporelles : les siècles sont composés de décennies qui sont elles-mêmes composées d'années. . . Les siècles sont eux aussi inclus dans une structure plus grande et ainsi de suite.

De plus le temps peut être exprimé à partir de différentes références. Il peut soit être exprimé de manière relative : pour des quantités « j'arrive dans 30 secondes » ou encore « j'ai 25 ans ». Il peut aussi être exprimé de manière absolue, pour des dates « je suis né en 1989 », et des durées « le match dure 90 minutes ».

La question de la précision est ici délicate car en réalité l'opération qui consiste à dater une photographie est particulière et elle demande des connaissances expertes du domaine. Elle requiert de solides connaissances du sujet de la prise de vue afin d'obtenir une haute précision. Dans quelle mesure cette précision est elle possible ? décennie ? année ? . . .

14. Plus les événements sont lointains dans le temps plus ils s'inscrivent dans une période temporelle large au contraire plus ils sont proches plus la granularité est fine. Cette perception se retrouve dans l'étude des périodes géologiques. Voir en annexe A.1 page 157 et les échelles temporelles présentées ici : <https://criticalape.wordpress.com/2011/09/29/comparing-time-scales-of-recorded-history-2/> dernier accès août 2015

1.2.1 Hypothèses du lien entre le texte et l'image

Dans cette thèse, nous abordons la question de la datation des images en nous intéressant au cas où l'image est accompagnée d'un texte qui est en rapport avec l'image. En effet, bien souvent, et notamment dans le web, une photographie va servir d'illustration et de complément au texte mais parfois c'est au contraire le texte qui décrit la photographie et permet sa compréhension. Il y a donc un lien plus ou moins fort entre le visuel d'une photographie et le contenu du texte qui peut se trouver à son côté. La sortie des systèmes que nous proposons dans cette thèse est une prédiction de la date de prise de vue de la photographie concernée. Ce que nous symbolisons ainsi :

$$f_{\text{datation}}(\text{image}, \text{texte}) = \text{date de prise de vue estimée}$$

où la fonction de datation f_{datation} prend en entrée une image et/ou un texte, et en sortie une date de prise de vue estimée par le système.

Selon nous, une photographie porte une information de temporalité dans son aspect et son contenu mais aussi dans son contexte environnant car :

- une photographie a un rôle informatif de par son contenu mais aussi temporel de par son état, son existence ;
- une photographie est par essence un témoignage du passé ;
- puisque toutes les photographies existantes appartiennent au passé, elles possèdent toutes une information temporelle inhérente, propre à la période de prise de vue.

C'est pourquoi nous avons émis l'hypothèse qu'il était possible de retrouver cette information temporelle et donc cette date.

Cependant, *peut-t-on dater toutes les photographies* ? Il existe des contenus *apriori* fortement temporels : par exemple, les bases d'images d'agences de presse. Selon nous, les photographies de presse ont pour vocation de marquer un événement. Nous retrouvons ici l'un des préceptes de base de la photographie : sauvegarder un instant selon une volonté de mémoire et d'archivage des données. À l'inverse, il existe des contenus dits atemporels : il s'agit ici de contenus ayant peu de lien avec le temps. Par exemple, des contenus artistiques, des photographies de détails dans laquelle on perd toute information de spatialité car celle-ci aide et est en lien avec la temporalité. Comme une photographie de désert : « Comment dater une photographie du désert de *Las Vegas* avant la construction des infrastructures ? ».

D'autres questions sont importantes à considérer. Par exemple, dans un texte, ce n'est pas parce que l'on parle d'un sujet ancien, *e.g.* une « traction », que la photographie est de l'époque pendant laquelle la traction a été populaire. En revanche cela nous indique *a minima*, que l'élément existe et est antérieur à l'écrit le mentionnant. Par exemple, si l'on parle du général de Gaulle dans un texte, il était nécessairement né au moment de sa rédaction, ce qui implique que ce texte est postérieur à sa date de naissance. De plus, dans la plupart des cas, il est très vraisemblable que l'image soit en lien avec la période évoquée dans le texte.

1.2.2 Dimensions de l'information visuelle

Une des premières applications de la datation automatique d'images est la classification d'images, qui consiste à attribuer une/des classe(s) ou étiquettes à une image. Un terme a souvent plusieurs dimensions d'interprétations possibles car il dispose d'une richesse et d'une variété thématique, géographique, culturelle, temporelle, *etc.* Par exemple, le simple terme « jaguar » peut faire référence à des concepts visuellement différents ayant chacun des dimensions thématiques associées (marque de voiture anglaise, tanks de l'armée allemande, félin, *etc.*).

Un moyen mis à notre disposition pour accéder à l'information visuelle et sa diversité sont les moteurs de recherche. Nous pouvons voir un exemple de résultat de requête textuelle en recherche d'images dans l'illustration 1.3. On y observe une classification visuelle thématique mettant en avant les plus globales, dans notre exemple : l'animal et les modèles récents de la marque de voiture.

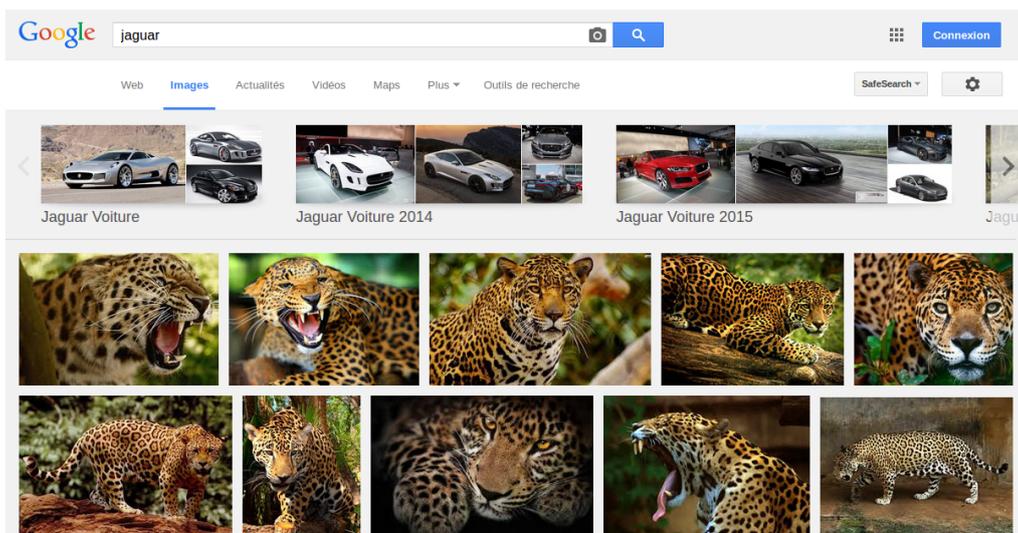


ILLUSTRATION 1.3: Les meilleurs résultats de recherche sont principalement focalisés sur la ou les dimensions thématiques prédominantes. (source Google images août 2015)

1.2. Datation d'images et applications

En effet, pour un sujet donné, la dimension thématique dépend de l'intention qui lui est associée. Tenir compte de la variété revient à tenir compte de l'intention de la requête. Par exemple si nous recherchons « petit déjeuner » nous obtenons des résultats illustrant le petit déjeuner européen/français car avec une telle requête le système suppose que nous sommes intéressés par le petit déjeuner français, il n'y a pas d'extrapolation vers le concept de petit déjeuner à travers les cultures et les pays. C'est pourquoi, si nous changeons de langue et que nous comparons les résultats visuels des requêtes effectuées¹⁵ nous observons une adaptation géographique et culturelle du concept du « petit déjeuner ».



ILLUSTRATION 1.4: Les particularités géographiques et culturelles sont déduites et contraintes par la langue de la requête. (source Google images août 2015)

C'est ce que nous pouvons observer dans l'illustration 1.4 : les résultats changent et tentent de correspondre au mieux à ma demande, l'intention de ma requête étant partiellement extrapolée par la langue de saisie. D'une manière générale, ces résultats sont surtout focalisés sur la culture et la localisation, qu'il est plus difficile de faire varier.

L'information temporelle, qui est l'aspect des images sur lequel nous nous sommes concentrés durant cette thèse, est une autre forme de variété selon laquelle classifier et indexer des bases d'images. Être capable de prédire la date d'une image ouvre ainsi la porte à de nouvelles applications en terme de recherche d'image, que nous présentons ci-dessous.

15. Toutes ces requêtes d'exemples ont été effectuées depuis le même ordinateur durant un bref laps de temps.

1.2.3 Recherche d'information temporelle

L'accès à l'information visuelle et temporelle peut se faire par les moteurs de recherche d'images grâce au texte ou à l'image¹⁶. Les résultats sont pour l'un, basés uniquement sur le contenu textuel et pour l'autre, basés sur le contenu visuel. Dans le premier cas, la précision de la réponse dépend donc de la justesse des mots-clés utilisés, dans le second cas, les réponses seront des images similaires à l'image utilisée comme requête.

Cependant, et c'est sur ce point que nous interviendrons, ces systèmes donnent des résultats ordonnés selon la proximité de la requête effectuée mais ils ne prennent pas en compte l'aspect temporel de celle-ci. Ceci se traduit dans les images résultantes de la recherche. Nous voulons donc trouver un moyen d'attribuer à chaque photographie la date de sa prise de vue ou au moins une estimation afin que par la suite la diversité temporelle des sujets soit considérée.

Prenons quelques exemples de requêtes pour lesquelles la temporalité est mal considérée dans les résultats de recherche d'images.

Si l'on recherche une personnalité vivante, d'une manière générale nous allons obtenir des photographies récentes plutôt qu'un panel représentatif de la vie ou de la carrière du sujet. Certains cas exceptionnels sont les célébrités ayant été connues et médiatisées dès leur plus jeune âge (Justin Bieber, Macaulay Culkin, Daniel Radcliffe, *etc.*).

Si l'on recherche un pays ou une ville, les résultats affichent de « belles » photographies qui sont certes utiles pour les utilisateurs (diaporama, fond d'écran, illustrations, *etc.*), mais qui occultent les faits temporels même proches dans le temps associés au lieu recherché. Par exemple, nous sommes aujourd'hui en 2015, et rechercher « Londres » ne donne plus prioritairement les résultats d'il y a 3 ans qui tous étaient afférents aux jeux olympiques de 2012 organisés dans cette ville.

Si nous recherchons un événement comme la « coupe du monde de football », les résultats sont généraux : le logo, la future ville et finalement peu de temporalité sur les villes historiques.

Nous observons que la variété temporelle est peu représentée. La temporalité de ces sujets est en fait surtout concentrée sur les faits récents plutôt que sur le caractère historique des sujets recherchés. D'une manière générale, les moteurs de recherche privilégient la ou les thématiques les plus communes et n'offrent pas de diversité temporelle riche.

Les moteurs de recherche prennent en compte la dénotation, *i.e.* le sens littéral représentatif d'un terme. Notre objectif est de prendre en compte la connotation et ainsi enrichir ces représentations avec, dans notre cas, la variété temporelle associable aux termes de recherches.

16. *e.g.* www.google.com/images ou encore www.tineye.com dernier accès août 2015

1.2.4 Exemples d'usage de l'information temporelle

L'enrichissement de photographies par leur dimension temporelle aura de multiples incidences sur l'utilisation quotidienne de moteurs de recherche et de nombreuses autres applications.

Désambiguïisation temporelle

Cela permet par exemple la désambiguïisation temporelle entre des homonymes existant à des époques temporelles différentes. Par exemple : les Ronaldo footballeurs, celui du Brésil et celui du Real de Madrid correspondent à deux époques différentes doivent par conséquent être désambiguïsés temporellement ; ou encore, les célébrités et leurs enfants (e.g. Bedos, Castaldi, pères et fils, etc.)

Visualisation de résultats de requêtes

Cela rend également possible des outils de présentation et de visualisation de résultats de requêtes, présentant l'évolution visuelle d'un monument, d'une ville, d'un événement ou une personnalité, etc. . Ainsi la visualisation, par exemple sous forme de frise chronologique permettra d'enrichir les résultats des moteurs de recherche et offrira une meilleure visualisation de l'information à l'utilisateur. Par exemple, le besoin d'information « *Je veux des photographies de la ville de New York* » exprimé alors avec la requête « New York » dans un moteur de recherche d'images par un utilisateur pourrait être satisfait par une réponse telle que dans l'illustration 1.5, qui permet d'observer l'évolution de ce sujet au cours du temps.



ILLUSTRATION 1.5: Exemple de résultats de recherche pour la requête « New York ».

Un tel système permet également de répondre à la demande d'un utilisateur exprimant une information simultanément thématique et explicitement temporelle, comme par exemple « *Je veux des photographies de Paris entre les années 1910 et 1930* ».

Continuité temporelle d'un sujet

De tels systèmes peuvent aussi assurer visuellement la continuité temporelle d'un sujet, comme proposé sur une personne dans l'illustration 1.6. Ce problème semble difficile mais une solution serait possible grâce à l'usage de l'information temporelle comme le montre l'illustration 1.7.

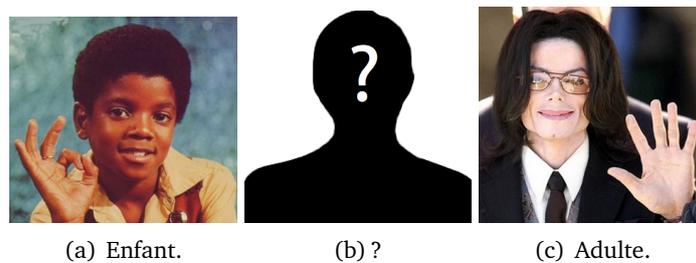


ILLUSTRATION 1.6: La continuité temporelle d'un sujet, ici une personnalité : Michael Jackson.



ILLUSTRATION 1.7: Exemple d'application possible, l'évolution temporelle d'un sujet, ici une personnalité : Michael Jackson [1958, 2009].

Si l'on utilise uniquement le contenu visuel des images, les possibilités s'étendent au-delà du web. Une telle approche peut par exemple être utilisée pour organiser des collections d'images personnelles ou bien stockées dans des archives (par exemple, de bibliothèques ou de musées).

Réorganisation d'albums ou de collections de photographies

Exploiter la temporalité permet aussi la réorganisation d'albums ou de collections de photographies voir l'illustration 1.8. Supposons que nous disposons de toutes les photographies numérisées appartenant à une famille et que nous décidons de les trier, nous pouvons imaginer un assistant qui ne daterait pas les photographies de manière absolue mais les organiserait de manière relative entre elles : telle personne est présente dans ces deux clichés et a vieilli alors cette photographie a été prise avant celle-ci. . .

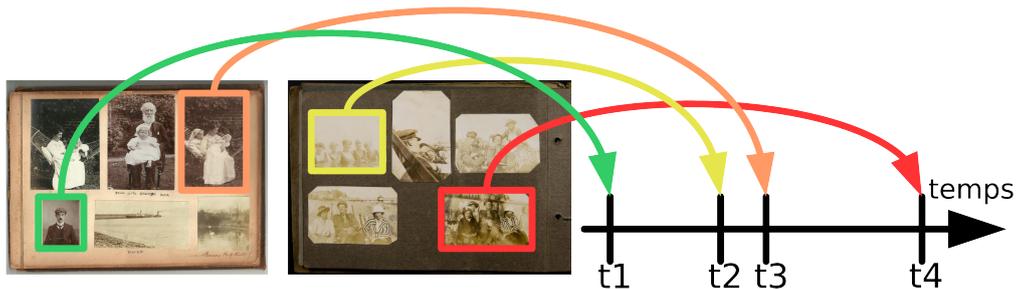


ILLUSTRATION 1.8: Exemple d'application possible, la réorganisation temporelle d'un album photographique numérisé. Ici $t1 < t2 < t3 < t4$ sont des marqueurs temporels non identifiés mais ordonnés.

Une variante de cette dernière application consiste en un assistantat, voire un remplacement complet de l'annotation manuelle par des experts du domaine pour aider les historiens dans la tâche de datation de photographies. D'autres applications sont encore possibles, comme par exemple la détection de faux ou la généalogie, etc.

1.3 Contributions et plan du manuscrit

Comme nous l'avons expliqué au cours de cette introduction, la problématique générale que nous abordons dans cette thèse est celle de la datation automatique des photographies anciennes à partir d'attributs issus de l'image et du texte qui l'entoure. Dans cette thèse nos travaux ont permis d'apporter un certain nombre de contributions, lesquelles font l'objet du reste de ce manuscrit.

Au cours de ces derniers siècles, les supports photographiques et les données associées¹⁷ ont subi de nombreuses mutations. Nous développerons dans le chapitre 2 nos observations au regard de ces avancées technologiques. Ce chapitre a une vocation informative et contextualisante. Il permet de comprendre l'évolution des méthodes de datation de photographies, que nous détaillons au chapitre 3.

17. Ces données associées sont définies au cours de cette thèse comme métadonnées.

Les méthodes de datation ont été historiquement manuelles avant de devenir récemment automatiques. En effet, ce problème a été introduit, simultanément dans les communautés de la vision par ordinateur (PALERMO, HAYS et EFROS, 2012) et de la recherche d'information (DIAS, MORENO, JATOWT et CAMPOS, 2012). Cependant, nous présentons aussi des travaux connexes, qui ont pour vocation d'inspirer les futures méthodes de datation.

Le chapitre 4, quant à lui, s'intéresse à une notion fondamentale de la datation : le caractère ordinal du temps. Nous présentons dans ce chapitre une synthèse des principales méthodes de classification ordinale, les mesures d'évaluation associées ainsi qu'un sous-ensemble des jeux de données couramment utilisés en classification ordinale.

Le chapitre 5 présente des jeux de données que nous avons construits pendant cette thèse. Nous avons en particulier conçu une large collection de pages web associant photographies anciennes et contenu textuel. Ces collections ont pour vocation de permettre à la fois l'entraînement et l'évaluation des modèles proposés. Elles représentent une avancée dans le domaine de la datation des photographies car ces dernières n'étaient pas disponibles en libre accès auparavant.

Le chapitre 6 est consacré à la datation des images à partir de leur contenu. Cette datation repose sur un modèle original que nous avons développé et évalué sur des bases d'images. La particularité majeure de notre approche est la mise en évidence le caractère ordinal du temps. Ceci constitue à notre connaissance la première utilisation de cette caractéristique temporelle intrinsèque à la datation de photographies.

Le chapitre 7 explore la datation des photographies par l'usage des données textuelles. Dans notre démarche, nous imitons le comportement humain utilisé lors de la datation de photographies insérées dans un contexte. Dans un premier temps, nous élevons à un niveau sémantique le contexte textuel grâce à l'exploitation des entités nommées. Ces dernières nous permettent d'obtenir des indices temporels que nous exploitons afin de déterminer une période de validité. Dans un second temps, il nous est ensuite possible de déterminer la date de prise de vue de la photographie. Ainsi nous proposons pour la première fois une méthode pour effectuer la datation des photographies par l'usage de l'environnement textuel.

Chapitre 2

Historique de la photographie et des métadonnées

Sommaire

2.1 Photographie	21
2.2 Métadonnées	25
2.3 Observations	30

Ce chapitre décrit très brièvement l'histoire de la photographie et des métadonnées. Nous nous sommes intéressés plus particulièrement aux méthodes de création de métadonnées associées aux différentes techniques photographiques. Nous passons alors en revue les méthodes traditionnelles (*e.g.* par annotation manuelle) et modernes (*e.g.* lors de la prise de vue par un appareil, ou dans le cadre d'un processus de numérisation).

2.1 Photographie

Nous abordons ici les différentes formes de photographies, dans un premier temps nous présentons brièvement les techniques traditionnelles historiques puis nous introduisons les techniques numériques et nous mettons en avant le changement de paradigme occasionné par cette métamorphose de la photographie. Nous avons répertorié certaines dates clés représentatives des évolutions majeures de la photographie dans une frise chronologique qui est consultable en annexe A.2 à la page 158.

2.1.1 Photographie traditionnelle

La photographie est le processus d'acquisition de scènes réelles par l'action de la lumière. À l'origine de la photographie, les usages étaient très limités en raison de contraintes techniques dues aux appareils et aux coûts engendrés. La photographie était un moyen mis en concurrence avec les arts plus traditionnels (peinture, sculpture...) et l'usage qui s'en dégageait relevait d'une volonté de mémoire. On se photographiait pour la postérité, à titre exceptionnel, événementiel, occasionnel. Traditionnellement, la captation de l'image a longtemps été réalisée sur un support matériel. Ce support physique a évolué à travers le temps dans le but d'améliorer et de simplifier les étapes liées au développement. Les matériaux les plus connus sont : les plaques de verre et de fer, les films souples ou encore le papier... Ces évolutions ont largement contribué à la popularisation de la photographie à travers le temps. Ainsi, la simplification et la miniaturisation des appareils ont permis aux photographes d'immortaliser quasiment toutes les situations. Du fait de sa simplification, la photographie s'est démocratisée au fil des années et est devenue, aujourd'hui, une activité « tout public ». Cette quête de la simplification a été initiée dès les débuts de la photographie : avec le slogan « vous appuyez sur le bouton, nous faisons le reste »¹, George Eastman (fondateur de *Kodak*) a commercialisé le premier appareil photographique portable, dès 1888 c.f. illustration 2.1. La conséquence de cette accessibilité à la technologie photographique se ressent au travers des usages et des activités. Les pratiques des utilisateurs se sont élargies au fil du temps et couvrent aujourd'hui quasiment tous les domaines : sur terre, sous l'eau, dans le cosmos et à toutes les échelles, aussi bien dans le microscopique que l'infiniment grand. La photographie est omniprésente : elle témoigne du passé, retrace l'histoire et perçoit le futur.

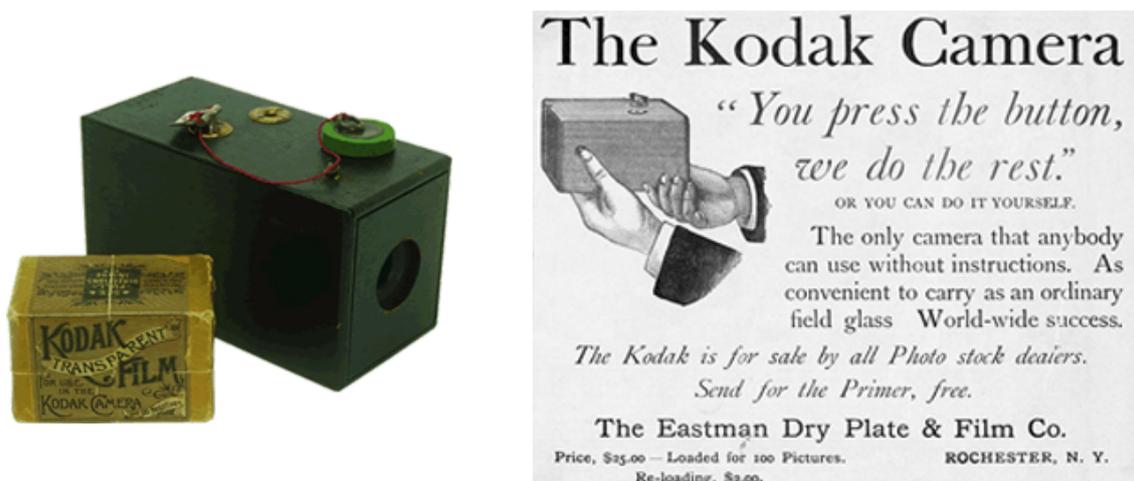


ILLUSTRATION 2.1: Publicité de 1889 pour le Kodak No.1 créé en 1888. (source Kodak)

1. « You press the button we do the rest » (publicité Kodak, 1889).

2.1.2 Photographie digitale

Depuis quelques décennies², la photographie a connu un nouvel essor dû à l'avènement des technologies modernes et de la numérisation. Les supports traditionnels (*i.e.* plaque de verre et de fer, papier, pellicule, *etc.*) ont laissé peu à peu la place au support numérique du fait de ses nombreux avantages, que nous développons ci-après. L'entrée des sociétés modernes dans l'ère digitale a provoqué l'adoption de la photographie numérique par le grand public : transfert via les réseaux de télécommunication, stockage de collection dans un faible volume physique, facilité de consultation des clichés, conservation dans le temps assurée par la non-altération des données digitales, duplication à coût zéro des données digitales, édition et retouches à l'infini via de nombreux logiciels *etc.* La photographie se dématérialise.

Impact du numérique

La photographie numérique a déclenché une véritable révolution, *c.f.* Annexe B. Elle s'est ressentie dans le système économique bâti autour de l'argentique, impactant le secteur d'activité lié à la photographie par la fermeture ou la reconversion de nombreuses entreprises impliquées dans la confection des consommables. L'entreprise *Kodak* a, par exemple, déposé le bilan en 2012, *c.f.* Annexe B.3. En effet, la photographie traditionnelle, et en particulier avec la technologie argentique, imposait au photographe une dépendance et une vigilance vis-à-vis des consommables engagés tels que la pellicule, le papier photographique, parfois les flashes à usage unique ou encore la pile. La photographie numérique quant à elle, au contraire, aurait tendance à déresponsabiliser le photographe puisqu'elle n'engendre maintenant plus d'impact financier direct. En effet, avec la technologie numérique, la photographie possède désormais une dimension immatérielle lui offrant l'intégralité de ces avantages à l'issue de la prise de vue. Grâce au numérique, il n'est plus nécessaire de développer les clichés. Ceci est un net avantage face à la norme établie par la technologie de l'argentique souvent associée à une longue et fastidieuse attente (*i.e.* envoi des pellicules, temps de développement, *etc.*) mais aussi à des ratés de développement liés à la configuration de l'appareil lors de la prise de vue. Le numérique offre un avantage économique considérable de ce point de vue. Il n'est donc pas nécessaire de développer les clichés pour les consulter, les partager, les stocker, les archiver ou encore pour les accumuler. Ces nouvelles pratiques ont métamorphosé l'usage de la photographie, produisant une nouvelle définition de celle-ci ; nous assistons, en effet, à une remise en cause de la dynamique présente depuis les origines de la photographie.

Il est donc indéniable que le numérique a littéralement redéfini les codes ancestraux de la photographie.

2. Premier appareil photographique numérique et premiers scanners à plat dans les années 1970, et boom de la photographie numérique à la fin des années 1990.

Numérisation des anciens clichés

Les procédés de digitalisation permettent également de numériser d'anciens clichés : soit par l'usage d'un scanner³, en numérisant le support sur lequel l'image a été inscrite (le négatif, le positif...), soit en prenant directement une photographie numérique du support matériel, c.f. l'illustration 2.2.



ILLUSTRATION 2.2: Le recto d'un ferrotype digitalisé dans son ensemble par un appareil photographique numérique. (source flickr « The Commons »)

L'inconvénient de la numérisation des clichés anciens réside dans le fait qu'une grande partie de l'information liée au cliché est perdue. Cette perte concerne les métadonnées. Une épreuve dépourvue de métadonnées conserve essentiellement une dimension esthétique. En dehors de cette dimension, un cliché perd du sens, voire la motivation de sa réalisation. Les métadonnées font partie intégrante de l'information constituant l'image. C'est pourquoi il est nécessaire de les préserver, ou de les (ré-)obtenir, afin de conserver le contexte d'interprétation du cliché.

3. C'est d'ailleurs ainsi qu'a été réalisée la première photographie digitale, visible dans l'illustration A.5 en annexe page 160.

2.2 Métadonnées

Nous définissons les métadonnées comme tous les éléments de contexte complémentaires qui sont propres à la prise de vue, mais qui n'en font pas rigoureusement partie. Nous détaillons par la suite les métadonnées traditionnelles et les métadonnées digitales. Pour certains clichés, la qualité de la composition de l'image permet de comprendre la majeure partie de l'information qui a été capturée. Mais, dans de nombreux cas et dans un souci de contextualisation, des informations devront être apportées pour compléter le discours entamé par l'image. Le caractère ponctuel et exclusif de la photographie entraîne un besoin de contextualisation par l'usage de notes, d'informations diverses, d'organisation de collection, de tri, de création d'albums, *etc.* Cela se traduit donc naturellement par une nécessité d'associer à tout élément photographique des données annexes permettant la compréhension de son contenu grâce à la contextualisation de celui-ci.

2.2.1 Métadonnées traditionnelles

La photographie est la captation d'un instantané, un arrêt sur image dans le flot du temps. Il est donc délicat de capter l'intégralité d'une situation par le viseur d'un appareil photographique. La photographie traduit une situation bien réelle en une représentation plus ou moins fidèle. À l'issue du développement, il est nécessaire de trier et organiser les photographies. L'archivage et le tri ont toujours été une tâche compliquée et fastidieuse, mais nécessaire. La mémorisation des conditions de réalisation des clichés l'est tout autant. Les photographes ont sans cesse été confrontés à ces problématiques, c'est pourquoi, dès les débuts de la photographie les métadonnées ont fait leur apparition. La première manifestation de la forte dépendance de la photographie aux métadonnées est le film *Kodak Autographique*. Ce film, voir l'illustration 2.3, été associé à un stylet en acier qui permettait d'inscrire, au dos des photographies au moment de leur réalisation, les informations de date, de lieu de prise de vue, de sujet *etc.* Des mentions textuelles sont donc venues très rapidement accompagner les photographies. Ces données contextualisantes ont facilité ainsi l'organisation et la compréhension du contenu des photographies.

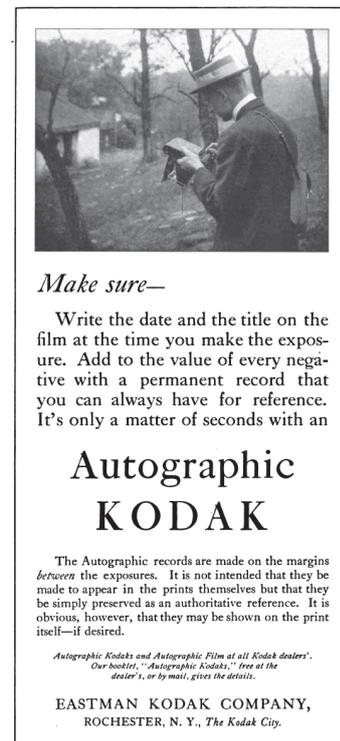


ILLUSTRATION 2.3: Publicité de 1915 pour le film Kodak Autographique. (source Kodak)

Métadonnées gérées manuellement à l'époque de l'argentique

Avec l'argentique, lorsque l'on réalisait des photographies, on annotait parfois des informations liées aux clichés réalisés sur des supports annexes. Ceux-ci servaient de légende afin de faciliter le classement, le tri mais aussi pour mémoriser le contexte des prises de vues, donnant un sens aux clichés. En raison de l'absence de normalisation, à l'époque de l'argentique, la diversité et les expressions des métadonnées s'illustraient selon différentes méthodes en suivant des procédés empiriques :

- notées sur la cassette de la pellicule : la sensibilité du film, la date, le nombre de vues, la marque, le type de film (couleur ou noir et blanc), la date et l'événement photographié ;
- inscrites automatiquement lors du développement au recto ou au verso du support : la date de développement, la marque du papier, la signature du photographe, le studio ayant développé l'épreuve, *etc.* ;
- dépendantes du support : le matériau du support (plaque de verre, daguerréotype, film souple, papier, *etc.*), le type et les finitions du papier utilisé (mat, brillant, bords blancs...), le grammage, *etc.* ;
- écrites au recto, au verso ou encore au côté de la photographie dans des albums, *etc.* ;
- stockées dans un carnet de notes tenu à jour lors de la réalisation des poses pouvant contenir les lieux et les paramètres de prises de vue tels que le type d'objectif utilisé ou encore les réglages du boîtier ou encore les conditions météo. . .

Arrivée des systèmes d'horodatage automatique des photographies



(a) Horodatage photographie argentique
10 octobre 1991.

(b) Horodatage photographie numérique
12 septembre ou 9 décembre 2009 ?

ILLUSTRATION 2.4: Exemples de différents horodatages et formats de dates selon les technologies photographiques.

Ce n'est que bien plus tard, dans les années 1990, que les métadonnées temporelles se sont inscrites automatiquement sur la face avant des photographies argentiques (*c.f.* illustration 2.4(a)), notamment grâce aux systèmes *Nikon Quartz Date Databack*, *Canon quartz date back E* ou encore au plus complexe *CHINON : INFO BACK-2 (DP-520) et INFO BACK (DB-010)*, qui permettaient quant à eux d'ajouter jusqu'à 30 caractères alphanumériques sur la photographie.

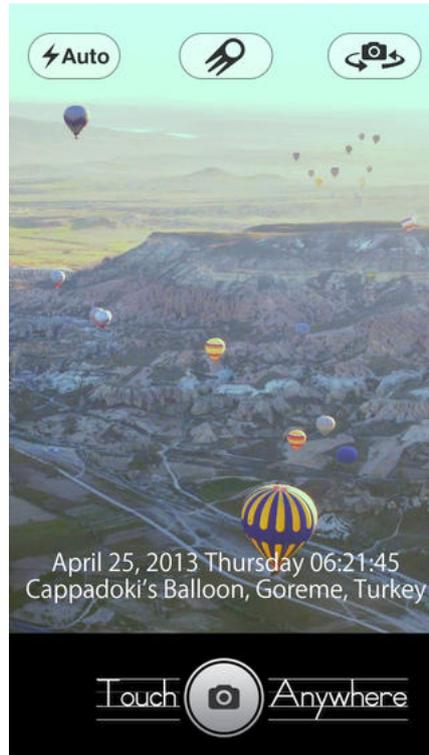


ILLUSTRATION 2.5: Ajout d'horodatages et de géolocalisation sur les images avec une application pour smartphone. (source « TimeStamp Photo with GPS Address Location » sur itunes)

Étrangement, bien que les données de date de prise de vue soient incluses dans les métadonnées numériques des fichiers générés, les appareils photographiques numériques ont eux aussi proposé cette possibilité d'inscrire la date sur l'image (c.f. illustration 2.4(b)). Il existe d'ailleurs toujours des applications (*Google Play/iTunes*⁴) pour smartphones permettant d'ajouter un horodatage, mais aussi des coordonnées géographiques, aux images (c.f. illustration 2.5).

2.2.2 Métadonnées digitales

Avec l'arrivée du numérique les métadonnées ont pris une autre dimension : elles sont venues surenchérir les métadonnées historiques de l'argentique. Le numérique a permis d'accéder à des informations plus précises, plus justes, et beaucoup plus complètes que celles offertes par l'argentique. Ainsi, grâce au numérique, les paramètres de l'appareil photographique peuvent être connus dans les moindres détails (c.f. illustration 2.6), ce qui tend à remplacer en partie l'usage des carnets de notes photographiques.

4. Ces deux exemples d'applications pour smartphone sont accessibles aux adresses suivantes : <https://play.google.com/store/apps/details?id=com.teapps.photostamperfree> et <https://itunes.apple.com/us/app/timestamp-photo-gps-address/id582538300?mt=8> dernier accès août 2015

Étiquette	Valeur
▼ Appareil photo	
Make	NIKON CORPORATION
Model	NIKON D5100
XResolution	300
YResolution	300
ResolutionUnit	pouces
Software	Ver.1.00
XResolution	300
YResolution	300
ResolutionUnit	pouces
ExifVersion	Version d'Exif inconnue
ComponentsConfiguration	Y Cb Cr -
CompressedBitsPerPixel	4
FocalLengthIn35mmFilm	52
InteroperabilityIndex	R98
InteroperabilityVersion	0100
▼ Données de l'image	
Orientation	En haut à gauche
DateTime	2011:06:25 18:31:01
YCbCrPositioning	Co-sited
Compression	Compression JPEG
YCbCrPositioning	Co-sited
ISOSpeedRatings	100
DateTimeOriginal	2011:06:25 18:31:01
DateTimeDigitized	2011:06:25 18:31:01
UserComment	

ILLUSTRATION 2.6: Exemple de données EXIF capturées avec un appareil photographique numérique.

Nous pouvons citer différents standards destinés à stocker les métadonnées de photographies digitales : *International Press and Telecommunications Council (IPTC)/ Information Interchange Model (IPTC/IIM⁵)*, *Exchangeable Image File Format (EXIF⁶)*, *eXtensible Metadata Platform XMP⁷*, qui intègre les données EXIF et IPTC/IMM, etc. Il existe de nombreux « standards » gérant les métadonnées digitales parmi lesquels nous pouvons également citer Geocodage, DIG35, JPX,

Peu à peu, les métadonnées ont été enrichies par des données GPS⁸ ou encore des données sonores⁹. D'une manière générale nous pouvons remarquer qu'il y a de plus en plus de métadonnées associées aux photographies. On peut d'ailleurs imaginer que les données météorologiques (pression atmosphérique, ensoleillement, force du vent, température, etc.), que les photographes avaient tendance à stocker dans leurs carnets de notes seront bientôt définies dans les métadonnées et automatiquement ajoutées lors de la prise de vue. En attendant, il existe déjà une méthode (ZHANG, KORAYEM, GRANDALL et LEBUHN, 2012) visant à déduire les données météorologiques des photographies à partir du contenu de l'image.

5. IIM : première version en 1991

6. EXIF : première version en octobre 1995

7. XMP : première version en avril 2001

8. Arrivée des modules GPS autour de 2008

9. *Audiosnaps* est un réseau social de partage de photographies permettant d'ajouter du son à une image avant de la partager.

D'autres méthodes encore visent à déduire des métadonnées telles que les coordonnées GPS à partir d'une prise de vue (HAYS et EFROS, 2008) ou encore ayant comme objectif de reconnaître un quartier d'une grande ville (DOERSCH, SINGH, GUPTA, SIVIC et EFROS, 2012) ou des sites touristiques (GAMMETER, BOSSARD, QUACK et GOOL, 2009)...

Cependant, bien que les métadonnées permettent de stocker des informations de temporalité comme des dates, le problème de la datation des œuvres anciennes persiste. En effet, plusieurs standards ont été développés. D'après les spécifications XMP d'Adobe nous pouvons lister et donner les définitions résumées de ces attributs de métadonnées temporelles :

`dc:date` Date et heure à laquelle quelque chose d'intéressant est arrivé à la ressource ;
`xmp:CreateDate` Date et heure de création originale ;
`xmp:MetadataDate` Date et heure du dernier changement du fichier ;
`xmp:ModifyDate` Date et heure de la dernière modification du fichier ;
`photoshop:DateCreated` Date et heure à laquelle le contenu intellectuel du document a été créé ;
`tiff:DateTime` Date et heure de la création de l'image ;
`exif:DateTimeOriginal` Date et heure de la génération de l'image originale ;
`exif:DateTimeDigitized` Date et heure du stockage de l'image comme donnée digitale ;
`exif:GPSTimeStamp` Horodatage provenant des données GPS exprimé en temps universel coordonné.

Un extrait des spécifications XMP d'Adobe concernant les données temporelles détaillées est disponible en annexe C.

Les nuances entre les définitions de ces informations de temporalité sont subtiles et parfois difficiles à interpréter. Néanmoins, d'après ces spécifications, les dates de prises de vue de photographies anciennes devraient se trouver dans l'attribut `photoshop:DateCreated`. Voici la description complète que nous trouvons concernant cette donnée dans la spécification XMP d'Adobe : « Il s'agit de la date à laquelle le contenu intellectuel du document a été créé (plutôt que la date de création de la représentation physique), d'après les conventions IIM. Par exemple, une photographie prise pendant la guerre civile américaine aurait une date de création au cours de cette époque (1861-1865) plutôt que la date à laquelle la photographie a été numérisée pour l'archivage. »¹⁰

10. Version originale : « *The date the intellectual content of the document was created (rather than the creation date of the physical representation), following IIM conventions. For example, a photo taken during the American Civil War would have a creation date during that epoch (1861-1865) rather than the date the photo was digitized for archiving.* ».

Cette information est bel et bien celle qui nous intéresse, et est catégorisée par *Adobe* en tant que métadonnée de catégorie « Externe ». À l'inverse des attributs de métadonnées de catégorie « Interne » qui sont renseignées automatiquement par le système de numérisation, ceux de catégorie « Externe » doivent être renseignés par l'utilisateur. Ces données dépendent donc de l'implication, de la précision et de la connaissance de l'utilisateur impliqué dans le processus de numérisation. Cela se traduit souvent par l'absence de ces données ou par une confiance en elles qui ne peut qu'être restreinte.

2.3 Observations

La datation des clichés nés à l'époque numérique peut sembler au premier abord être un faux problème car, comme vu précédemment, les données de dates de prises de vue sont techniquement incluses dans les métadonnées des photographies numériques. Bien que ces métadonnées soient théoriquement valides, il persiste des inconsistances du fait de l'erreur humaine. En effet, ces métadonnées générées par les appareils photographiques numériques dépendent de la bonne configuration de l'appareil par l'utilisateur. Il en résulte parfois des photographies antidatées. Nous remarquons ainsi par exemple une quantité non négligeable d'images datées du 1^{er} janvier 1970 ou encore du 1^{er} janvier 2000. Ce phénomène est dû aux dates par défaut utilisées dans les systèmes d'exploitation, le plus connu étant le *timestamp* UNIX¹¹. Cette problématique peut concerner tous les systèmes de numérisation.

Ce phénomène s'observe notamment avec les données en provenance de *flickr* lorsque l'on recherche sur le site¹² des photographies prises le 1^{er} janvier 1970, qui correspond dans le système du site à 0, nous observons une quantité importante de photographies numériques automatiquement mal horodatées.

En ce qui concerne les données que nous avons manipulées au cours de la création de notre jeu de données (voir le chapitre 5 à la page 71) nous reportons les statistiques liées aux métadonnées numériques que nous avons extraites dans l'annexe D, à la page 171.

11. Nombre de secondes depuis le 1^{er} janvier 1970 00 :00 :00 UTC

12. https://www.flickr.com/search/?text=&min_taken_date=0&max_taken_date=0 dernier accès août 2015

Chapitre 3

Datation de photographies : des méthodes manuelles aux automatiques

Sommaire

3.1 Origines de la datation de photographies : méthodes manuelles . . .	31
3.2 Méthodes automatiques de datation des photographies numérisées	35
3.3 Méthodes automatiques pour la datation des écrits	41
3.4 Propositions vis-à-vis des méthodes existantes	44

Nous présentons dans ce chapitre le lien historique entre leurs contenus textuels et visuels, mis en avant lors de la datation de photographies. C'est pourquoi, nous avons mené des recherches bibliographiques, à la fois, dans la communauté de la vision par ordinateur, mais aussi, dans celle du traitement automatique des langues. Ceci afin de profiter des expertises plus ou moins adaptées ou encore adaptables de ces communautés.

Nous avons recherché dans les littératures populaires, sociales et scientifiques comment le problème de la datation de photographies a-t-il été traité par le passé.

3.1 Origines de la datation de photographies : méthodes manuelles

Pour associer une période temporelle à une photographie, les méthodes traditionnelles utilisent à leur avantage le support physique du cliché mais aussi bon nombre d'indices présents en dehors du contenu visuel de la photographie. D'une manière générale ces méthodes exploitent l'historique de l'évolution des procédés, de leur commercialisation et de leur popularité.

3.1.1 Photographie à travers le temps

Comme indiqué dans le chapitre 2, la photographie est un domaine en perpétuelle évolution, tant par ses usages que par les raisons qui suscitent la réalisation de clichés. Ces différents usages accompagnent l'évolution des techniques photographiques à travers le temps. Il est important de considérer que l'évolution des usages de la photographie n'est pas continue et qu'elle dépend avant tout des changements des techniques d'acquisition de l'information visuelle. Nous pouvons citer comme exemple les écarts des fréquences d'usages. Aux origines de la photographie, les usages étaient exclusifs et ponctuels (e.g. souvenir ultime pour la postérité). Les tendances ont évolué jusqu'à aujourd'hui et les techniques se sont démocratisées. Les usages actuels sont à l'inverse beaucoup moins raisonnés et démesurément actifs (e.g. albums de photographies sur les réseaux sociaux).

L'origine et la motivation de la prise de vue influencent aussi la temporalité de celle-ci comme nous l'avions évoqué dans nos hypothèses à la sous-section 1.2.1. Il sera, de par leur nature, plus difficile de dater des photographies d'art que de presse. En effet, les photographies de presse s'inscrivent dans le présent lors de leur réalisation et elles ont souvent pour vocation de relater l'histoire à un moment donné. À l'inverse, les photographies d'art, pour lesquelles la notion de créativité est plus mise en avant, peuvent avoir une vocation abstraite qui est parfois dénuée de temporalité¹.

De plus, il est à noter que la technologie photographique utilisée a aussi un rôle dans cette étape de datation. Nous pouvons supposer qu'il est possible de dater une image à partir du type d'appareil photographique utilisé (e.g. déformation, résolution, aberrations chromatiques, couleurs). Selon cette hypothèse nous pourrions ainsi disposer d'une date « minimale » pour définir la date de prise de vue du cliché. Cette date serait la date de mise en service du type d'appareil (c.f. illustration 3.1) ou de la technologie photographique associée.



ILLUSTRATION 3.1: L'évolution des technologies photographiques. (source popchart-lab.com)

1. Dans l'art, la temporalité se manifeste sous forme de mouvements, de tendances, de courants et de modes qui sont dépendantes des artistes avant tout.

3.1.2 Démarche de l'expertise manuelle pour la datation de photographies

À présent nous allons répondre à la question suivante : « Quelles sont les techniques qui permettent de dater une photographie manuellement à partir de l'objet matériel et de son support physique ? » Les évolutions de la photographie dans ses usages, ses motivations et ses techniques sont autant de sources d'information pour la datation de clichés. Ainsi non seulement le contenu de l'image sert, mais aussi les informations externes : c'est ce qui nous inspire à tirer parti de l'image mais aussi du texte environnant.

Grâce à la technologie photographique utilisée

Les premières méthodes comptaient sur une analyse manuelle des propriétés physiques du média (e.g. le type de papier, la taille, les enduits, le type de couleurs, etc.). COE et HAWORTH-BOOTH (1983) proposent de dater manuellement les photographies en reconnaissant les techniques photographiques les plus anciennes.

Grâce à la technique photographique utilisée et au format

L'auteur, HARDING (*Blog du « National Media Museum » de Grande Bretagne*) donne son expertise pour apprendre à reconnaître différents types de photographies dans une série d'articles nommés « Dater vos anciennes photographies grâce au format et à la technique photographique utilisée ». L'auteur annonce qu'il est possible de dater une photographie sans considérer ce qui est photographié mais en considérant le type de photographie à laquelle nous avons à faire. Puis il décrit les procédés et les formats mis en place pendant les 100 premières années de la photographie commerciale.

Il explique notamment comment reconnaître les techniques photographiques suivantes :

- un daguerréotype (années 1840 – années 1850) ;
- un positif sur plaque de verre au collodion humide, aussi connu sous le nom d'ambrotype (début des années 1850 – années 1880) ;
- un positif sur plaque de fer au collodion humide, aussi connu sous les noms de ferrotypage et *tintype* ou encore *melainotype* (1855 – années 1940).

Il détaille aussi les formats caractéristiques de certaines photographies historiques et explique comment reconnaître :

- un format « carte de visite » (fin des années 1850 – c.1910) ;
- un format « carte cabinet » (1866 – c.1914) ;
- un format « carte postale » (1900 – années 1950).

Cette méthode part du principe que les procédés et les formats utilisés n'étaient populaires que pendant une période de temps limitée, ainsi cette méthode donne rapidement une approximation de la période de réalisation du cliché observé.

Concernant les formats et les dimensions voici ce que nous trouvons sur wikipédia² :

Les petites photographies, au format 5,2 cm sur 8,7 cm, sont ensuite collées sur :

- un carton qui adopte le format d'une « carte de visite », 6,2 cm sur 10,3 cm ;
- un format encore plus petit, appelé « mignonnette » (carton 4 cm sur 8 cm), mais qui ne connaîtra pas le même engouement ;
- un format plus grand, sur carton de 10,5 x 16,5 cm, qui est appelé « carte cabinet ».

Grâce au contenu

En revanche, si nous souhaitons dater une image à partir de son contenu (par opposition à son support), de nombreuses références offrent des outils pour aider la datation. Ces outils sont fondés sur les caractéristiques visuelles des objets, des lieux ou des personnes apparaissant dans les photographies.

En effet, le contenu donne une information sur la date potentielle de prise de vue. À ce sujet, il existe un nombre important de recueils pouvant tracer l'histoire d'un thème, comme par exemple l'évolution des modes (PEACOCK, 2008).

Pour citer quelques autres exemples accessibles en ligne, le site du « *Victoria and Albert Museum* » explique comment dater une image à partir des vêtements, sur la période de 1840 à 1930, alors que le site du « *English Museums Association* » propose une méthodologie d'évaluation complète comprenant quelques-uns des détails stéréotypiques suivants : les vêtements et les poses prises par des personnes, la composition de l'image, les scènes *etc.*

Grâce au contenu, au support et aux métadonnées.

« *Datation des anciennes photographies* » est une ressource intéressante sur la datation manuelle de photographies. Sur ce site, nous trouvons des éléments permettant la datation manuelle de photographies anciennes à partir du support réel de la photographie. Il offre une expertise sur les dates de mise en service des différents types de photographie.

Les points à considérer se trouvent d'une part dans les caractéristiques physiques du support. Il faudra pour cela observer notamment les faces avant et arrière, le type de matériaux, l'épaisseur, les dimensions mais aussi les finitions esthétiques du support.

D'autre part, il est important de considérer que des indices temporels se trouvent aussi dans le contenu de la photographie. En effet, les contraintes techniques liées à la photographie³, les coupes de cheveux, les accessoires, le décor utilisé, *etc.* influent sur le type ou la manière avec laquelle les clichés sont réalisés.

2. https://fr.wikipedia.org/wiki/Photo-carte_de_visite dernier accès août 2015

3. Les débuts de la technologie photographique imposaient une stature figée due au temps de pose nécessaire pour qu'une image nette soit capturée.

Enfin, les métadonnées servent aussi à la datation des clichés car elles donnent de précieux indices temporels : le nom, l'adresse du photographe ou du studio ou encore la marque de l'imprimerie qui a fait le développement.

Grâce au contenu et au support de manière semi-automatique

Selon les mêmes principes mais cette fois-ci d'une manière semi-automatisée le site « *DIY Photo Dating Wizard* » propose un « magicien » capable de trouver la date de prise de vue d'une photographie. Ce système est uniquement destiné aux portraits et aux photographies de familles anglaises des époques Victoriennes⁴ et Édouardiennes⁵ ayant des formats « mignonnette », « carte de visite » ou « carte cabinet ». En pratique, le système propose un ensemble de questions successives basées sur les caractéristiques du support et du contenu. Les réponses que nous fournissons fonctionnent comme pour le parcours d'un arbre de décision, amenant ainsi à une estimation de la date de prise de vue.

Cependant, bien qu'étant fonctionnelles, les méthodes mentionnées ci-dessus nécessitent une intervention humaine, et parfois même celle d'experts de la période concernée, et ne sont par conséquent pas applicables à grande échelle.

3.2 Méthodes automatiques de datation des photographies numériques

Les photographies numériques sont obtenues par une digitalisation qui peut être effectuée par le biais d'appareils photographiques numériques ou bien de scanners. Lors de la numérisation⁶, une grande partie de l'information du support original est perdue car seul le contenu de la prise de vue est capturé. Bien évidemment, une approche possible est d'utiliser l'horodatage numérique dans les métadonnées, mais cette information est disponible seulement pour un petit ensemble d'images récentes nées à l'époque numérique et dépendent de la configuration de l'appareil. Nous allons donc nous focaliser sur les méthodes automatiques, utilisant le contenu des pixels des images numériques, car ces méthodes sont utilisables à grande échelle.

3.2.1 Approches par détection de filigranes

Nous pouvons considérer une première famille d'approches fonctionnant par détection de filigranes et utilisant des systèmes de reconnaissance optique de caractères (ROC).

4. L'époque Victorienne désigne les années de règne de Victoria I^{re} au Royaume-Uni, de 1837 à 1901.

5. L'époque Édouardienne désigne les années de règne d'Édouard VII au Royaume-Uni, de 1901 à 1910.

6. Nous avons présenté la numérisation des anciens clichés dans la sous-section 2.1.2.

Une première méthode (CHEN et ZHANG, 2003) recherchant à dater des photographies numériques automatiquement est concentrée sur un type particulier d'images digitales. Les images pour lesquelles la date de prise de vue est inscrite en surimpression sur la face avant de la photographie, aussi appelées images avec horodatage en filigrane, *c.f.* illustrations 2.4 à la page 26. Ainsi, cette méthode vise à la fois à détecter mais aussi reconnaître l'information de temporalité inscrite dans les pixels des images digitales. La variété et la complexité de l'arrière-plan des filigranes rendent parfois leurs lectures difficiles, car l'arrière-plan influence la qualité de l'impression des filigranes (*e.g.* la saturation de l'impression). Malgré ces contraintes techniques, leur méthode atteint de très bons scores de détection et de reconnaissance.

Leur méthode de détection et de reconnaissance utilise une modélisation des horodatages et procède par recherche dans le contenu des photographies en limitant les possibilités grâce à l'*a priori* de localisation des horodatages spécifiques à cette tâche. CHEN et ZHANG (2003) ont modélisé différents formats d'écriture des dates répondant aux particularités choisies par les constructeurs (Mois - Jour - Année ou Année - Mois - Jour). Ces formats permettent de représenter de nombreux paramètres tels que la distance entre les lettres, la taille et l'angle de l'écriture, la position dans l'image, *etc.* Ils ont aussi modélisé les polices utilisées dans l'écriture des horodatages (police 7 segments et matrice de points). Leur approche utilise des déductions *a priori* de la zone de recherche, de la taille et l'angle d'écriture, *etc.* L'objectif est de réduire considérablement l'espace de recherche permettant ainsi d'augmenter la rapidité d'exécution de leur méthode.

Récemment, la société *Eastman Kodak*⁷ a développé un brevet (SCALISE et BLOSE, 2012) pour la datation automatique de photographies. Leur système utilise les faces avant et arrière des images pour en déduire une période de temps pendant laquelle le traitement de la pellicule aurait pu être fait.

Leur approche est décomposée en plusieurs phases. Tout d'abord, cette méthode vise à détecter les filigranes sur les deux faces numérisées d'une photographie. Les filigranes peuvent être des horodatages ou des logos en surimpression ou encore des indications se trouvant au dos des clichés tels que la marque et le type de papier utilisé, la date de développement, *etc.* *c.f.* illustration 3.2. Ensuite vient une phase visant à trouver la correspondance de ces filigranes, l'objectif étant de comparer ces filigranes à une base de données de filigranes pour lesquels les dates de mise en service sont bien connues. Enfin, à partir des informations temporelles recueillies, la période possible correspondant à la date d'élaboration du traitement photographique est déterminée.

7. Demande de brevet (SCALISE et BLOSE, 2010) et Octroi du brevet (SCALISE et BLOSE, 2012).



ILLUSTRATION 3.2: Différents filigranes se trouvant au verso d'impressions photographiques, variant selon les époques photographiques et les constructeurs. (source des images WEAVER et LONG (2009))

Cependant, le principal inconvénient de cette méthode est qu'elle soit concentrée sur certains cas très spécifiques. Le problème étant qu'il est rare que le verso de la photographie soit numérisée et mis à disposition en même temps que son recto. Par conséquent, le champ d'application de cette méthode est limité en termes de redocumentarisation de données volumineuses.

Remarque : certains musées comme le MoMA⁸ mettent à disposition sur leur site web une collection d'image (*Modern Photographs 1909–1949 : The Thomas Walther Collection*) offrant les deux faces numérisées ainsi que des informations extrêmement précises sur le support (épaisseur, poids, technique photographique utilisée, composition chimique du support analysé, etc.), la surface (réflexion des UVs, le contexte de l'image indiquant si elle fait partie d'une série ou non, etc.).

Nous pouvons citer deux travaux connexes exploitant les filigranes :

Premièrement, GUY (2008) mentionne dans un rapport une méthode automatique de détection de dates inscrites sur la face avant des photographies numérisées et de retouche par « *inpainting* » de la zone cible supprimant la mention de date de la zone considérée.

8. *The Museum of Modern Art*

Deuxièmement, WEYAND, TSAI et LEIBE (2015) proposent une solution au problème des faux positifs retrouvés lors de la recherche d'images similaires contenant des filigranes, des horodatages ou des cadres de décorations. Le problème étant que les images paraissent similaires⁹ car elles disposent d'une même représentation des zones surimprimées. La méthode propose de détecter les marques et permet de réaliser le calcul de similarité et d'association des images sans considérer les zones « communes » surimprimées. Cette méthode permet ainsi d'éviter les faux positifs.

3.2.2 Approches utilisant des caractéristiques visuelles

Dans leurs travaux, DIAS, MORENO, JATOWT et CAMPOS (2012) proposent de s'attaquer au problème de la datation de photographies grâce à des caractéristiques extraites des images. Étant donné la nouveauté de cette problématique, ils ont alors créé leur jeu de données. Pour ce faire ils ont utilisé une technique d'extension temporelle de requêtes textuelles (LOUPY, BELLOT, EL-BÈZE et MARTEAU, 1998 ; XU et CROFT, 1996) basées sur des noms de villes. Ainsi, après filtrage, ils ont obtenu des images des villes aux périodes visées. Ils proposent ensuite d'utiliser une méthode automatique basée sur des caractéristiques visuelles afin de retrouver la période de prise de vue des clichés. À cet effet, ils utilisent 3 caractéristiques visuelles couleurs de l'état de l'art : ScalableColor (CIEPLINSKI, 2001), FCTH (CHATZICHRISTOFIS et BOUTALIS, 2008b) et CEDD (CHATZICHRISTOFIS et BOUTALIS, 2008a) dans une chaîne de traitement usuelle d'apprentissage supervisé avec un SVM¹⁰ (VAPNIK, 1995) linéaire standard avec les paramètres par défaut¹¹. Il est important de remarquer qu'ils proposent un découpage du temps déduit d'un *a priori* sur le contexte de cette tâche. En effet, les classes ne sont ni des années ni des décennies mais des périodes temporelles correspondant aux évolutions majeures de la photographie. Leur méthode est testée sur un jeu de données de 1 170 images de villes appartenant aux 5 périodes temporelles suivantes :

- 1826 – 1925 : Premières photographies ;
- 1925 – 1948 : Démocratisation des appareils noir et blanc ;
- 1948 – 1968 : Démocratisation des appareils couleurs ;
- 1968 – 1982 : Apparition des premiers appareils réflexes ;
- 1982 – 2011 : Appareils numériques (jusqu'à la date des expériences réalisées).

Ce problème de classification à 5 classes est évalué avec les métriques standards (BAEZA-YATES, RIBEIRO-NETO et al., 1999) suivantes : la précision, le rappel, la F-Mesure et l'aire sous la courbe ROC¹² et donne des résultats encourageants.

9. Du point de vue des algorithmes de recherche d'images similaires.

10. Séparateur à Vaste Marge

11. `C=1...` ils ont utilisé `libsvm` CHANG et LIN (2011).

12. *Receiver Operating Characteristic*

Durant la même année mais dans un domaine différent, PALERMO, HAYS et EFROS (2012) ont proposé une nouvelle approche permettant de prédire la décennie au cours de laquelle une photographie a été réalisée en se basant sur son contenu. Pour ce faire, ils ont constitué un jeu de données comportant 1 325¹³ photographies uniquement en couleur. Les dates de prises de vue de ces photographies s'étalent des années 1930 à la fin des années 1970. Ce jeu de données étalé sur 5 décennies est équilibré, ainsi, chacune des décennies comporte 265 clichés. Leur objectif est la classification de chacune de ces photographies dans l'une de ces 5 décennies. Dans leur approche, la contribution principale se concentre sur les caractéristiques visuelles utilisées plutôt que sur le modèle d'apprentissage. Le modèle utilisé est un SVM linéaire utilisant une stratégie *Un contre Un* classique en classification supervisée. Les caractéristiques extraites quant à elles sont principalement des informations de couleurs car leur but est de capturer une information temporellement discriminante basée sur l'évolution au fil du temps des processus photographiques couleurs. Ils utilisent en tout 7 caractéristiques d'images : en plus de trois caractéristiques d'images populaires en classification d'images (Descripteur GIST (OLIVA et TORRALBA, 2001), *Tiny images* (TORRALBA, FERGUS et FREEMAN, 2008) et des Histogrammes de couleurs L*a*b* (« Sun database : Largescale scene recognition from abbey to zoo »)), ils ont mis au point quatre nouvelles caractéristiques d'image spécifiques au problème de la datation de photographies :

- Caractéristique de similarité des processus des pellicules couleurs ;
- Histogrammes de cooccurrence couleur ;
- Probabilité conditionnelle de la saturation sachant la teinte ;
- Histogrammes de teintes.

Ces nouveaux descripteurs sont introduits et présentés en détails dans leur article (PALERMO, HAYS et EFROS, 2012). Ils effectuent une validation croisée à 10 plis et choisissent de manière aléatoire dans chaque pli 80% de la base pour l'apprentissage, soit 1 075 images, et 20% pour le test, soit 250 images. Ils évaluent ensuite leur classification à 5 classes avec la précision¹⁴ moyenne sur 10 tirages aléatoires. Les résultats de classification montrent une précision de 45,7% dans le meilleur des cas. À titre de comparaison, la précision d'un humain non entraîné sur leur base d'images est de 26% alors que celle d'une classification aléatoire est de 20%. Ceci est donc une approche intéressante car elle offre des résultats supérieurs à l'humain (+19,7 points).

13. Dans l'article, le nombre d'images de leur jeu de données est annoncé à 1 375 par erreur, nous reportons ici le nombre d'images présentes dans l'archive disponible à l'adresse <http://graphics.cs.cmu.edu/projects/historicalColor/> dernier accès août 2015

14. En anglais : *Accuracy*

Suite à cette approche, d'autres méthodes basées sur des caractéristiques de couleur traduisant l'évolution temporelle de la photographie ont été proposées, notamment par FERNANDO, MUSELET, KHAN et TUYTELAARS (2014). Les caractéristiques couleurs développées dans leurs travaux sont basées sur les angles que forment les couleurs des pixels¹⁵ une fois projetés dans un espace RVB¹⁶. En effet, les auteurs prouvent que les angles produits sont discriminants entre les types d'appareils photographiques utilisés. C'est pourquoi cette caractéristique, qui capte une information historique vis-à-vis de l'appareil utilisé, donne de meilleurs résultats que ceux de PALERMO, HAYS et EFROS (2012) sur le même jeu de données.

Nous pouvons remarquer que l'usage de descripteurs couleurs donne en effet de bons résultats sur cette base. Par exemple, il a été démontré par BERENQUER ARROYO et VIDAL (2013) que l'usage du descripteur *Color Naming* (BENAVENTE, VANRELL et BALDRICH, 2008) permet d'améliorer les résultats initiaux de PALERMO, HAYS et EFROS (2012). Le descripteur *Color Naming* dans cette approche peut être vu comme une discrétisation de l'espace des couleurs. Cette discrétisation transforme tous les pixels couleurs d'une image en un vecteur de 11 dimensions indiquant les 11 probabilités représentatives des 11 couleurs que l'être humain emploie le plus couramment¹⁷.

Il existe des travaux connexes qui s'intéressent à dater les styles visuels temporels, que ce soient par exemple des voitures (LEE, EFROS et HEBERT, 2013) ou des façades d'immeubles (LEE, MAISONNEUVE, CRANDALL, EFROS et SIVIC, 2015). Ces travaux visent à associer un objet contenu dans une image à la date de création de celui-ci. Les travaux de ARDIZZONE, DINDO et MAZZOLA (2009) quant à eux, s'intéressent à l'évaluation des types de dégradations que subissent les photographies. À l'inverse, il est aussi parfois intéressant d'ignorer les effets de la temporalité. REMATAS, FERNANDO, TOMMASI et TUYTELAARS (2013) considèrent la temporalité comme un domaine spécifique, au sein duquel les aspects visuels des objets évoluent et ils tirent parti de cette continuité pour améliorer la classification sans toutefois être impactés par l'évolution temporelle des classes. FERNANDO, TOMMASI et TUYTELAARS (2015) s'intéressent à comparer visuellement des images temporellement espacées, leur objectif est d'associer des images représentatives d'un même monument ensemble et ce quelque-soit la période de réalisation de la photographie. Ils proposent des caractéristiques visuelles qui sont peu affectées par les effets visuels imposés par les changements technologiques de la photographie. *e.g.* les caractéristiques développées permettent de considérer visuellement proche des photographies de la tour Eiffel du début du xx^e siècle avec des prises de vue modernes de la tour Eiffel.

15. Inspirés des techniques calibration radiométrique de LIN, GU, YAMAZAKI et SHUM (2004).

16. Rouge, Vert et Bleu : espace colorimétrique à 3 composantes.

17. Des études ont montré que l'humain utilise majoritairement 11 termes pour définir l'ensemble des couleurs (BOYNTON et OLSON, 1990 ; STURGES et WHITFIELD, 1997) il s'agit des 11 termes suivants : rouge, orange, marron, jaune, vert, bleu, violet, rose, noir, gris et blanc.

3.3 Méthodes automatiques pour la datation des écrits

À notre connaissance, aucun travail ne traite du problème de la datation de photographies par l'usage de caractéristiques textuelles. À l'inverse, il existe une méthode (HE, SAMMARA, BURGERS et SCHOMAKER, 2014) visant à dater les textes en utilisant des techniques de vision par ordinateur. Ils y évaluent l'évolution des styles d'écriture grâce à l'usage de caractéristiques utilisées dans la caractérisation des auteurs associées à des techniques de régression globales et locales. Cependant, il existe des travaux traitant de la datation des textes (anciens et nouveaux) et pour lesquels des analogies sont possibles avec la datation des images.

Parmi ces travaux, deux grandes familles peuvent être définies : l'une cherchant à dater l'époque durant laquelle l'écrit a été réalisé et l'autre visant à dater la période concernée par l'écrit. Les travaux de JONG, RODE et HIEMSTRA (2005) ; KANHABUA et NØRVÅG (2008) visent à dater la création du document en lui attribuant un horodatage¹⁸. A *contrario*, les travaux de JATOWT, AU YEUNG et TANAKA (2013) visent à retrouver la temporalité du contenu du document.

3.3.1 Approches visant à retrouver la date de création de l'œuvre

Les premiers travaux proposent un modèle statistique du langage¹⁹ temporel (JONG, RODE et HIEMSTRA, 2005) permettant de dater un document texte²⁰. Des perfectionnements de ces travaux ont été réalisés par KANHABUA et NØRVÅG (2008). Leurs améliorations portent sur deux parties importantes de cette méthode.

Premièrement, ils valorisent les données d'entrée du système en se basant sur des approches sémantiques. Pour ce faire, ils utilisent notamment un étiquetage morpho-syntaxique (BRILL, 1993 ; SCHMID, 1994) permettant ainsi de sélectionner seulement les mots les plus importants en se basant sur les étiquettes obtenues. Mais aussi, ils exploitent l'information apportée par l'extraction de collocations (MANNING et SCHÜTZE, 1999) afin de garantir la bonne compréhension du sens de certains termes. Pour améliorer l'interprétation des mots, ils ont fait appel à des méthodes de désambiguïsation lexicale (NAVIGLI, 2009). L'extraction de concepts leur permet de compenser la rareté des termes de certains documents. Enfin, ils effectuent le filtrage des mots les moins informatifs grâce aux valeurs de *tf-idf* (RAMOS, 1999).

18. En anglais : *Timestamp*

19. Les modèles statistiques du langage ont été largement utilisés dans la traduction automatique (BROWN, COCKE, PIETRA, PIETRA, JELINEK, LAFFERTY, MERCER et ROOSSIN, 1990).

20. Des modèles temporels du langage ont aussi été proposés pour répondre aux problématiques de recherche d'information temporelle BERBERICH, BEDATHUR, ALONSO et WEIKUM (2010).

Deuxièmement, ils proposent des améliorations du modèle statistique du langage temporel (JONG, RODE et HIEMSTRA, 2005). *Primo*, ils conçoivent un processus d'interpolation temporelle des mots avec deux stratégies : une pour les mots qui se répètent dans le temps et une autre pour les mots plutôt « diffus » qui ont une incidence temporelle ponctuelle. *Secundo*, ils utilisent un système d'entropie temporelle (LOCHBAUM et STREETER, 1989) visant à mesurer l'importance d'un mot au sein de la collection. Leur but étant de mettre en avant les termes représentant au mieux les périodes temporelles. *Tertio*, ils apportent des statistiques externes en complément comme par exemple des statistiques de recherche de *Google Zeitgeist* avec la granularité des semaines.

3.3.2 Approches visant à comprendre la période concernée par l'œuvre

Heideltime (STRÖTGEN et GERTZ, 2010) est un système pour l'extraction et la normalisation des expressions temporelles dans les textes. L'extraction et la normalisation²¹ sont basées sur des règles : motifs d'expression régulières et étiquetage morpho-syntaxique du voisinage (BRILL, 1993 ; SCHMID, 1994). Cependant, les ressources doivent contenir des mentions d'informations temporelles explicites pour être détectées. Une fois ces informations temporelles détectées il est possible de les manipuler. STRÖTGEN, ALONSO et GERTZ (2012) ont proposé d'identifier les dates les plus importantes dans un document compte tenu des positions et des fréquences de celles-ci dans les phrases. Ils ont aussi proposé d'élire les dates les plus importantes dans un document en fonction d'une requête textuelle associée.

JATOWT, AU YEUNG et TANAKA (2013) introduisent la notion de temporalité du contenu du document²² qui est opposée en tant qu'information complémentaire à la date de création du document²³. Ils définissent et catégorisent premièrement ce que sont les documents temporels, qu'ils opposent aux documents atemporels et deuxièmement ce que représente la temporalité au sein d'un document. Pour déterminer la temporalité d'un document temporel, ils utilisent des bases de connaissances externes issues d'une collection composée de nouvelles publiées dans des articles de journaux pour lesquelles des informations temporelles sont explicitement disponibles.

Leur approche est divisée en 3 phases.

21. Les informations temporelles annotées par *Heideltime* suivent les spécifications données par TimeML (PUSTEJOVSKY, CASTAÑO, INGRIA, SAURÍ, GAIZUSKAS, SETZER et KATZ, 2003).

22. En anglais : *Document focus time*

23. En anglais : *Timestamp*

Première phase : ils calculent l'association temporelle des mots contenus dans leur base de connaissance. Pour ce faire, ils créent un graphe d'association entre les mots (technique notamment utilisées avec des entités nommées (NADEAU et SEKINE, 2007) dans les travaux de GROSS, DOUCET et TOIVONEN (2013)). Ils considèrent les mentions de dates comme des mots, ce qui leur permet alors de relier des mots avec des périodes de temps spécifiques.

Deuxième phase : ils déterminent les mots qui seront les plus discriminants temporellement, et qui par conséquent doivent être considérés avec une attention particulière. Ils attribuent des poids selon l'importance de la relation entre un mot et la période temporelle dans laquelle il apparaît. Plus un mot est typique d'une époque plus il sera discriminant pour cette époque et donc recevra un poids important.

Troisième et dernière phase : à partir des informations d'associations temporelles des mots et des poids décrivant la capacité discriminante des mots, les auteurs sont capables de déterminer la temporalité d'un document. Ils proposent une approche intuitive qui suppose qu'un document appartient à l'époque qui est la plus représentée par les termes les plus discriminants qui le composent. À partir de cette « intuition », ils ont développé d'autres méthodes.

Récemment des travaux se sont concentrés sur le lien entre les entités nommées et les photographies historiques. ZHANG, JATOWT et TANAKA (« Search for Images of Historical Objects Using Wikipedia ») ont abordé le problème de la recherche d'images historiques représentatives d'entités nommées par l'usage de *Wikipédia*.

De nouvelles approches (DIAS, HASANUZZAMAN, FERRARI et MATHET, 2014) visent à retrouver la temporalité des termes inscrits dans *WordNet* (MILLER, 1995). Grâce au *TempoWordNet* il est ainsi possible de connaître la polarité temporelle d'un terme, c'est à dire, savoir s'il est temporel ou au contraire atemporel. Lorsqu'un terme est temporel il est possible d'obtenir son appartenance au *continuum* temporel, c'est à dire, au passé, au présent ou au futur.

3.4 Propositions vis-à-vis des méthodes existantes

Nous nous sommes posés les questions suivantes : La temporalité doit-elle être considérée à part entière?, La datation de photographies est-elle un problème de classification standard ?

Les approches sur les caractéristiques visuelles ne prennent pas en compte l'aspect temporel (FERNANDO, MUSELET, KHAN et TUYTELAARS, 2014 ; DIAS, MORENO, JATOWT et CAMPOS, 2012 ; PALERMO, HAYS et EFROS, 2012). En effet, la datation est traitée comme un problème standard de classification multi-classes. Cependant, dans le cadre d'une estimation de la date de prise de vue d'une photographie, les classes étudiées suivent un ordre chronologique. Il serait donc intéressant d'en tenir compte dans les modèles d'apprentissage (FRANK et HALL, 2001), et pour le moins, dans les mesures d'évaluation de leurs performances. Ceci dans le but d'observer pour les erreurs commises les distances avec les valeurs réelles. Avec des mesures de ce type, il est tout à fait probable que l'humain soit plus performant car, même s'il ne dispose pas de l'information juste, son évaluation est une estimation sans doute proche de la réalité. À l'inverse, les classifieurs donnant de mauvaises réponses n'ont aucune raison de donner des résultats proches des valeurs réelles car il opposent lors de la phase d'apprentissage toutes les classes entre elles sans même considérer l'ordre temporel qui les relie. En effet, pour les mesures d'évaluation classiques, la différence entre la date réelle et la date prédite n'est pas prise en compte. Se tromper d'une année ou d'une décennie revient à la même performance, alors que ce ne devrait pas être le cas.

Nous avons montré que la datation relève en réalité des problèmes de classification ordinaire. Les méthodes de classification ordinaire ont déjà été utilisées en vision par ordinateur pour répondre à des problèmes d'estimation d'âge de personnes (CHANG, CHEN et HUNG, 2011) ou de taille de véhicules (SHASHUA et LEVIN, 2003) à partir de photographies.

Nous introduisons les méthodes de classification ordinaire dans le chapitre 4 puis nous présentons dans le chapitre 6 une méthode de classification ordinaire que nous avons développée et appliquée à la datation de photographies.

Du point de vue du texte plusieurs aspects sont à considérer. Tout d'abord, aucun jeu de données n'est disponible pour évaluer la datation des images en utilisant les textes les contextualisant. Nous avons ainsi créé un jeu de données, présenté au chapitre 5. Pour traiter ce problème, nous développons une approche de la temporalité par bases de connaissances, à l'aide de connaissances temporelles externes sur des entités nommées (NADEAU et SEKINE, 2007 ; RITTER, CLARK, ETZIONI et al., 2011), que nous décrivons dans le chapitre 7.

Chapitre 4

Classification ordinale

Sommaire

4.1 Fondements théoriques	46
4.2 Définitions	49
4.3 Méthodes d'évaluations de la classification ordinale	56
4.4 Jeux de données ordinaux	65
4.5 Discussion	69

La classification ordinale (FRANK et HALL, 2001 ; WAEGEMAN et BOULLART, 2006) parfois appelée régression ordinale (LI et LIN, 2007 ; HERBRICH, GRAEPEL et OBERMAYER, 1999), « *ranking* », « *ranking learning* » ou encore « *learning to rank* » est un genre particulier de classification dans lequel les classes sont organisées suivant une relation d'ordre. Cette caractéristique implique des spécialisations, aussi bien dans la conception des modèles d'apprentissage que dans leur évaluation. Toutes ces considérations sont fondées sur le fait que l'ordre intrinsèque des données doit être appris et évalué de manière appropriée.

La classification ordinale est par nature plus proche du comportement humain qu'une classification binaire. En effet, un jugement humain aura plutôt tendance à s'exprimer selon une échelle ordinale allant d'un extrême à l'autre plutôt que de façon strictement binaire. C'est pourquoi de nombreuses méthodes existantes dans la communauté d'apprentissage automatique ont été adaptées pour tenter de répondre à cette problématique qui propose d'ordonner les données. De ce fait, de nombreux ensembles de données ordinaux sont basés sur l'évaluation de produits et de services par des utilisateurs. Ces évaluations se traduisent en effet naturellement de façon ordinale car elles s'expriment sur une échelle graduée dans la majorité des systèmes.

La classification ordinaire peut être définie comme une tâche hybride qui partage des caractéristiques provenant à la fois de la classification et de la régression. En effet, comme dans la classification, les classes sont discrètes, mais comme dans la régression il y a un ordre entre elles. Toutefois, HERBRICH, GRAEPEL, OBERMAYER et al. (1999) précisent que (traduction condensée) : « contrairement à la régression, (...) l'espace des étiquettes ne définit pas un espace métrique ; la distance entre les étiquettes des éléments n'est pas définie, seul l'ordre entre les étiquettes a un sens ». Par conséquent, nous préférons utiliser l'appellation classification ordinaire plutôt que régression ordinaire.

Dans les sections suivantes, nous allons présenter les méthodes utilisées pour résoudre les problèmes de classification ordinaire, les mesures d'évaluation et une liste de jeux de données existants associés à ces contextes.

4.1 Fondements théoriques

L'apprentissage statistique supervisé, est basé sur l'utilisation d'exemples d'apprentissages $\{(x_1, y_1), \dots, (x_n, y_n)\}$ pour lesquels nous voulons apprendre un modèle appelé : fonction hypothèse $h : X \rightarrow Y$ qui, à partir d'instances x_1, \dots, x_n retourne des étiquettes y_1, \dots, y_n . Dans cette représentation, $X \subseteq \mathbb{R}^d$ est l'espace de départ aussi appelé espace des caractéristiques et Y l'espace d'arrivée aussi appelé espace des étiquettes. Afin d'obtenir le modèle répondant le mieux à notre problématique, il est nécessaire au cours de l'apprentissage de minimiser un coût attendu (noté L), entre les valeurs prédites par celui-ci et les valeurs réelles. Ce coût est associé à un risque appelé risque fonctionnel, noté R (VAPNIK, 1999). De tels problèmes peuvent donc être formulés en tant qu'estimation de la Minimisation du Risque Empirique (noté ERM suivant l'acronyme anglais) :

Définition 1. *Minimisation du Risque Empirique (ERM) :*

Le risque fonctionnel attendu R est remplacé par le risque empirique, ce qui produit :

$$R_{emp}(h) = \frac{1}{n} \sum_{i=1}^n L(h(x_i), y_i) \quad (4.1)$$

Nous recherchons donc la meilleure fonction hypothèse h , que nous noterons \hat{h} , qui minimise le coût associé L :

$$\hat{h} = \operatorname{argmin}_{h \in H} R_{emp}(h) \quad (4.2)$$

Nous pouvons à présent définir la fonction de coût qui permettra la calibration du modèle et ainsi obtenir la meilleure fonction hypothèse \hat{h} .

Définition 2. *Fonction de coût (L) ou d'optimisation :*

Une fonction de coût est utilisée pour comparer les valeurs prédites par le modèle avec les valeurs originales observées, aussi appelées vérité terrain.

La conception de cette fonction dépend de l'objectif que nous voulons faire apprendre au modèle.

Dans les sous-sections suivantes nous allons faire un rapide tour d'horizon des différentes méthodes d'apprentissage statistique supervisé en lien avec la classification ordinale ainsi que des fonctions de coûts associées à ces méthodes. Pour plus de détails sur ces techniques, voir (VAPNIK, 1999 ; BISHOP, 2006 ; CHEN, LIU, LAN, MA et LI, 2009).

4.1.1 Classification

Traditionnellement dans la classification multi-classes, les étiquettes Y n'ont pas de lien entre elles et constituent un ensemble discret fini de possibilités. Il n'y a ni ordre, ni métrique entre les étiquettes. Ainsi, si les étiquettes sont différentes, les instances de classes et les classes associées sont intégralement distinctes, sans aucune forme de pondération.

C'est pourquoi les erreurs sont traitées sans distinction comme par exemple par la fonction de coût très usitée, L_{0-1} :

$$L_{0-1}(h(x), y) = I(h(x) \neq y) = \begin{cases} 1 & \text{si } h(x) \neq y \\ 0 & \text{si } h(x) = y \end{cases} \quad (4.3)$$

où $I \in [0, 1]$ est la fonction indicative qui pénalise l'erreur d'association entre les valeurs prédites et les valeurs réelles. Avec une telle fonction, toutes les erreurs sont pénalisées de la même manière.

4.1.2 Régression

Contrairement à la classification, en régression, l'espace des étiquettes Y représente l'ensemble des valeurs réelles. Il y a donc un ordre entre les étiquettes possibles. De plus, l'espace Y est un espace métrique, ce qui veut dire qu'employer des notions de distance entre les instances a un sens. Les fonctions de coût utilisées dans la régression sont basées sur des notions de distance entre les valeurs prédites et les valeurs réelles.

Il est possible de choisir de nombreuses fonctions de coût, parmi lesquelles deux sont très fréquemment utilisées pour la régression, $L_{absolue}$ et $L_{carré}$.

La fonction $L_{absolue}$ calcule la distance absolue entre l'étiquette prédite et l'étiquette réelle :

$$L_{absolue}(h(x), y) = |h(x) - y| \quad (4.4)$$

La fonction $L_{carré}$, quant à elle, calcule la distance élevée au carré entre l'étiquette prédite et l'étiquette réelle :

$$L_{carré}(h(x), y) = (h(x) - y)^2 \quad (4.5)$$

Nous pouvons remarquer que la fonction de coût $L_{carré}$ pénalise très fortement les instances marginales¹ du fait de l'élévation au carré de la distance.

4.1.3 Classement

Nous nous concentrons uniquement sur le classement² supervisé pour lequel un ensemble de préférences entre les instances $P = \{x_1 < x_2, x_3 < x_2, x_4 < x_3 \dots\}$ est donné.

Cet ensemble donne des informations sur les préférences relatives, par paire³, entre les instances. Ces préférences sont déduites de connaissances préalables. Elles peuvent venir de critères discriminants, un critère discriminant étant un critère pour lequel des mesures de similarité ont un sens vis-à-vis de l'objectif défini par le classement général. Le but du classement est donc de trouver l'ordre qui réponde le mieux à l'objectif visé pour un ensemble global d'instances. La difficulté est que cet ordre global n'est pas prédéfini et que seuls les ordres locaux peuvent être définis. Parmi l'ensemble des ordres locaux prédéfinis, toutes les paires d'instances sont classées soit supérieures, inférieures ou égales entre elles.

La famille des fonctions de coût utilisées dans le classement est basée sur des comparaisons par paire : l'ordre des instances x_i et x_j , équivalent à leur rang, est ainsi donné par :

$$L_{rang}(h(x_i), h(x_j), y_i, y_j) = \begin{cases} 1 & \text{si } y_i > y_j \\ & \text{et } \neg h(x_i) > h(x_j) \\ 1 & \text{si } y_j > y_i \\ & \text{et } \neg h(x_j) > h(x_i) \\ 0 & \text{sinon} \end{cases} \quad (4.6)$$

1. Les instances marginales sont les instances dont les prédictions sont les plus distantes de leurs valeurs attendues.

2. Souvent dénommée par le terme anglais : *Ranking*

3. En anglais : *Pairwise approach*

Cette fonction de coût a pour but de pénaliser les paires, dont les valeurs prédites et les valeurs réelles sont discordantes.

Nous allons maintenant introduire la classification ordinale au regard des méthodes que nous venons de présenter. La classification ordinale emprunte dans la définition de son modèle des éléments de la classification et de la régression. En revanche, les mesures d'évaluation qui lui sont associées sont inspirées des mesures utilisées dans la régression et dans le classement.

4.2 Définitions

Tout comme les méthodes de classification, de régression et de classement supervisé, les méthodes de classification ordinale font partie de la famille des méthodes d'apprentissage supervisé. Comme nous l'avons annoncé dans l'introduction, elles peuvent être définies comme des méthodes hybrides entre les méthodes de classification et de régression. Dans la classification, l'ensemble des classes est considéré discret et fini. Dans la régression, l'espace d'arrivée est ordonné. Ainsi, nous pouvons définir $Y = \{y^{(1)} < y^{(2)} < \dots < y^{(K)}\}$ en temps qu'ensemble discret, fini et ordonné de K étiquettes de classes. Contrairement à la régression, nous rappelons⁴ que l'espace Y n'est pas un espace métrique. La notion de distance entre les classes n'a donc pas de sens dans cet espace.

4.2.1 Approche générale

La différence principale avec la classification est qu'il existe, dans la classification ordinale, un ordre dans l'espace des étiquettes.

Nous voulons apprendre à partir de l'ensemble d'apprentissage $S_{\text{apprentissage}}$ une généralisation \hat{h} de la fonction hypothèse $h : X \rightarrow Y$ qui, à partir des caractéristiques x_j est capable de prédire \hat{y}_j la valeur ordinale.

Le modèle appris, appelé le classifieur, noté \hat{h} , va associer à chaque objet x_j l'une des K étiquettes dans Y . Notons donc $\hat{h}(x_j) = \hat{y}_j$, où \hat{y}_j représente la prédiction. Le but est d'associer au mieux les étiquettes aux caractéristiques extraites des données du problème.

Lorsque nous disons « au mieux », dans la classification ordinale, nous entendons respecter au maximum ces deux critères :

- D'une part, avec le plus de précision possible au sein de l'ordre défini par l'ensemble des étiquettes *i.e.* minimiser $D(\hat{y}_j, y_j)$ la distance entre la valeur prédite et la valeur réelle (à l'égal de la famille des fonctions de coûts utilisées dans la régression, *c.f.* sous-section 4.1.2) ;

4. Comme indiqué par HERBRICH, GRAEPEL, OBERMAYER et al. (1999).

- D'autre part, de manière consistante pour un ensemble de plusieurs prédictions. Autrement dit, pour un ensemble de données de test, noté $S_{test} = \{(x_1, y_1), \dots, (x_i, y_i), \dots, (x_n, y_n)\}$ où $y_1 \leq \dots \leq y_i \leq \dots \leq y_n$, nous voulons garder l'ensemble des valeurs prédites systématiquement ordonnées selon l'ordre présent dans S_{test} , sans toutefois considérer l'exactitude des valeurs prédites. Ainsi nous voulons autant que possible minimiser les inversions dans l'ordre : $\hat{y}_1 \leq \dots \leq \hat{y}_i \leq \dots \leq \hat{y}_n$ (à l'égal de la famille des fonctions de coûts utilisées dans le classement, *c.f.* sous-section 4.1.3).

4.2.2 Familles d'approches

Il existe de nombreuses méthodes d'apprentissage supervisé qui abordent le problème de la classification ordinaire. Ces méthodes peuvent être organisées en branches ou familles.

Les deux grandes familles sont les approches fonctionnant par paire et les approches qui, après projection, fonctionnent par seuil.

Les approches par paire (HERBRICH, GRAEPEL, OBERMAYER et al., 1999), comme leur nom l'indique, traitent le problème par évaluation sur les paires d'instances des différentes classes. Ainsi, les ordres entre les paires d'instances sont comparés. L'ordre entre les valeurs réelles des deux instances est comparé à l'ordre entre les valeurs prédites de ces deux mêmes instances. Cet ordre relatif donne une information de concordance ou de discordance entre les valeurs réelles et les valeurs prédites. Seul cet ordre relatif est évalué et la prédiction absolue n'est pas visée. Ces approches visent au contraire à respecter un classement ordinal plutôt que la prédiction des valeurs associées aux instances.

A contrario, les approches par seuil (CRAMMER et SINGER, 2001 ; SHASHUA et LEVIN, 2003 ; LIN, ABU-MOSTAFA, ENGINEERING et SCIENCE, 2008) visent à projeter les données sur un axe décrivant la nature ordinaire du problème. L'objectif étant ensuite de définir des délimitations sur cet axe sous forme de seuils permettant le découpage le plus juste possible, selon des classes ordonnées. Alors, les prédictions peuvent être effectuées en fonction de la projection des instances sur cet axe échelonné.

Il existe également d'autres approches plus difficiles à catégoriser, comme par exemple :

- celles qui effectuent des traitements de quantification sur les étiquettes (KRAMER, WIDMER, PFAHRINGER et GROEVE, 2001) ;
- celles qui considèrent le problème de classification ordinaire comme un problème de classification monotone (POTHARST et BIOCH, 1999) ;
- celles qui visent à découper le problème principal en sous-problèmes binaires (FRANK et HALL, 2001 ; WAEGEMAN et BOULLART, 2006 ; CARDOSO et COSTA, 2007) ;
- *etc.*

Par la suite, nous allons nous concentrer sur les méthodes qui fonctionnent par décomposition en sous-problèmes binaires. Ces méthodes sont faciles d'accès, elles ne requièrent pas de changements majeurs des algorithmes utilisés, et donnent, comparativement aux autres méthodes de classification ordinale, de bons résultats. Nous allons à présent développer et expliciter leurs fonctionnements.

4.2.3 Approche par décomposition en sous-problèmes binaires

Nous allons d'abord introduire les méthodes non ordinales permettant la décomposition en sous-problèmes binaires afin de traiter des problèmes multi-classes⁵, puis nous présenterons la stratégie ordinale.

Les approches par décomposition en sous-problèmes binaires ont été utilisées par le passé car elles étendent les capacités des classifieurs binaires à traiter les problèmes de classification multi-classes (BISHOP, 2006).

Le fondement de ces approches réside dans 3 phases.

La première phase concerne la stratégie de ré-étiquetage des classes : Les classes de Y sont ré-étiquetées en *negatives* et *positives* selon la stratégie choisie pour l'apprentissage des sous-classifieurs. Nous pouvons citer les deux stratégies les plus répandues pour la décomposition en sous-problèmes binaires : la stratégie *Un contre le Reste*⁶ (VAPNIK, 1995) et la stratégie *Un contre Un*⁷ (KNERR, PERSONNAZ et DREYFUS, 1990). Nous allons comparer ces stratégies avant d'introduire leur pendant en classification ordinale qui a été proposé par FRANK et HALL (2001) que nous nommons en temps que stratégie *Avant contre Après*.

La deuxième phase est l'apprentissage de l'ensemble des sous-classifieurs selon les classes ré-étiquetées.

La troisième phase est la dernière étape qui permet la prédiction de la classe de d'appartenance de l'instance testée. Les sous-classifieurs appris sont utilisés pour la prédiction. À ce stade, nous montrerons les différents *scenarii* qui ont été proposés dans la littérature pour la stratégie ordinale *Avant contre Après*.

5. C'est à dire lorsque $K > 2$.

6. En anglais : *One vs Rest*

7. En anglais : *One vs One*

Notations

Afin d'illustrer ces concepts théoriques nous allons supposer que nous nous plaçons dans le cadre d'un problème ordinal de datation de photographies que nous définissons comme suit. L'objectif est de prédire au mieux la date de prise de vue de photographies. Soit $S_{\text{apprentissage}} = \{(x_1, y_1), \dots, (x_i, y_i), \dots, (x_n, y_n)\}$ la collection d'apprentissage composée de n photographies datées, dans laquelle $x_i \in X \subseteq \mathbb{R}^d$ représente les caractéristiques extraites d'une photographie et $y_i \in Y$ est l'une des K dates possibles (e.g. une année, une décennie...).

Le problème consiste donc à prédire, de la meilleure manière possible, la date de prise de vue, notée \hat{y}_j , d'une nouvelle photographie x_j .

Nous voulons donc apprendre à partir de l'ensemble des photographies datées connues $S_{\text{apprentissage}}$ une généralisation \hat{h} de la fonction hypothèse $h : X \rightarrow Y$ qui, à partir des caractéristiques des photographies est capable de prédire les dates de prises de vue de celles-ci.

Nous proposons de dater les photographies grâce à 5 étiquettes qui représentent 5 classes ordinales comme ci-dessous :

- y_1 : correspond aux photographies prises au cours de la décennie 1930 ;
- y_2 : correspond aux photographies prises au cours de la décennie 1940 ;
- y_3 : correspond aux photographies prises au cours de la décennie 1950 ;
- y_4 : correspond aux photographies prises au cours de la décennie 1960 ;
- y_5 : correspond aux photographies prises au cours de la décennie 1970.

Dans un premier temps, nous pouvons donc considérer ce problème comme un problème de classification multi-classes, en ne prenant pas en compte le caractère ordinal induit par les étiquettes. Alors, comme énoncé précédemment, parmi les stratégies les plus répandues (BISHOP, 2006) pour effectuer une classification multi-classes à partir de classifieurs binaires, nous pouvons choisir entre les stratégies suivantes :

- *Un contre le Reste* (VAPNIK, 1995) ;
- *Un contre Un* (KNERR, PERSONNAZ et DREYFUS, 1990).

Afin d'explicitier le fonctionnement de ces différentes stratégies de ré-étiquetage d'instances nous utilisons les notations suivantes : C_i représente toutes les instances de la classe y_i . Le but d'un classifieur binaire est de différencier les instances *negatives* des instances *positives*. La sortie de ce classifieur binaire est donc l'attribution de nouvelles instances à l'une de ces 2 classes. C'est pourquoi nous noterons $Cl = (\textit{negatives}, \textit{positives})$ le sous-classifieur Cl appris à partir de l'opposition des instances *negatives* aux instances *positives*.

Un contre le Reste

La stratégie de ré-étiquetage *Un contre le Reste*, parfois aussi appelée *Un contre Tous*⁸, produira autant de sous-problèmes binaires que de classes présentes dans le problème (K). Chaque sous-problème est appris à partir des instances d'une classe (e.g. C_1) étiquetées comme *negatives* et le reste des instances des autres classes (e.g. $C_{2...5}$) étiquetées comme *positives*. Ainsi, chacune des classes est mise à l'écart pour être opposée aux autres, dans chacun des sous-classifieurs. Ceci produit par conséquent 5 sous-classifieurs binaires répartis comme suit :

$$\begin{aligned} S_{Un\ contre\ le\ Reste} &= \{Cl_1, Cl_2, Cl_3, Cl_4, Cl_5\} \\ &= \{(C_1, \{C_2, C_3, C_4, C_5\}), \\ &\quad (C_2, \{C_1, C_3, C_4, C_5\}), \\ &\quad (C_3, \{C_1, C_2, C_4, C_5\}), \\ &\quad (C_4, \{C_1, C_2, C_3, C_5\}), \\ &\quad (C_5, \{C_1, C_2, C_3, C_4\})\} \end{aligned}$$

$S_{Un\ contre\ le\ Reste}$ est l'ensemble de tous les classifieurs appris avec la stratégie *Un contre le Reste*.

Les avantages de cette stratégie sont, d'une part qu'elle permet d'utiliser les données de toutes les classes en même temps lors de la phase d'apprentissage, et d'autre part, que peu de classifieurs sont nécessaires. En effet, avec cette stratégie, un problème à K classes ne requiert que K classifieurs⁹.

Une des faiblesses de cette stratégie provient du fait qu'il n'y a pas de phase de vote des différents classifieurs car c'est celui qui émet la prédiction avec la confiance maximale qui va déterminer la classe prédite. Un autre point faible concerne le déséquilibre des sous-problèmes appris. En effet, les sous-problèmes produits seront très fortement déséquilibrés entre la classe *negative* et la classe *positive*. Par exemple, si pour un problème à 5 classes, les classes d'origine sont équilibrées entre elles, chaque sous-problème opposera toutes les instances d'une classe donnée aux instances de toutes les autres classes, qui seront donc 4 fois plus nombreuses.

Un contre Un

La stratégie *Un contre Un* produira des classifieurs binaires opposant toutes les paires de classes possibles. Chaque sous-classifieur oppose ainsi toutes les instances d'une classe à

8. En anglais : *One vs All*

9. Nous sommes ici dans le cadre de la classification de multi-classes ($K > 2$). Lorsque $K = 2$ nous ne sommes pas dans une configuration multi-classes alors un seul classifieur binaire suffit.

toutes les instances d'une autre classe. Cette stratégie produit ainsi dans notre exemple $\binom{5}{2} = 10$ sous-problèmes binaires. Le premier classifieur noté Cl_1 opposera donc toutes les instances de la classe y_1 notées C_1 , étiquetée alors comme la classe *négative*, aux instances de la classe y_2 notées C_2 , étiquetée alors comme la classe *positive*. Les classifieurs suivants concernent à leur tour les autres classes du problème. Cette stratégie produit donc 10 classifieurs binaires répartis comme ceci :

$$\begin{aligned} S_{U_n \text{ contre } U_n} &= \{Cl_1, Cl_2, \dots, Cl_{10}\} \\ &= \{(C_1, C_2), (C_1, C_3), (C_1, C_4), (C_1, C_5), \\ &\quad (C_2, C_3), (C_2, C_4), (C_2, C_5), \\ &\quad (C_3, C_4), (C_3, C_5), \\ &\quad (C_4, C_5)\} \end{aligned}$$

$S_{U_n \text{ contre } U_n}$ représente l'ensemble de tous les classifieurs binaires appris qui opposent pour chaque paire toutes les instances de chacune des classes concernées.

Le vote est l'une des forces de cette stratégie car pour un problème à K classes, $\binom{K}{2} = K(K-1)/2$ classifieurs binaires sont appris et participeront au vote final pour chacune des K classes. Toutes les classes reçoivent ou non des votes de l'ensemble des classifieurs, ces votes sont utilisés pour élire la classe majoritaire.

Sa faiblesse est donc le nombre élevé de classifieurs qui doivent être entraînés, ce qui peut s'avérer problématique dans certains cas. L'un des autres inconvénients réside dans le fait que peu de données sont utilisées pour chacun des sous-classifieurs. En effet, ne sont utilisées pour l'apprentissage d'un sous-classifieur que les données de deux classes.

Avant contre Après

La stratégie, que nous nommons *Avant contre Après*, proposée par FRANK et HALL (2001) est un moyen simple d'exploiter les informations ordinales présentes dans les problèmes ordinaux. Le seul pré-requis est de disposer de classifieurs produisant des estimations de probabilité de classe sans qu'il soit nécessaire d'apporter de modifications à l'algorithme utilisé. Cette astuce transforme un problème ordinal à K classes en $K-1$ sous-problèmes de classification binaire.

Dans cette approche, l'ensemble des étiquettes Y est considéré comme un axe ordonné. Pour chacun des sous-classifieurs appris, cette stratégie va utiliser l'intégralité des classes en les opposant de part et d'autre d'une séparation de cet axe. Ainsi si nous considérons i comme étant un séparateur de l'ensemble ordinal, avec i allant de 1 à $K-1$, alors les classes ré-étiquetées comme *négatives* seront les classes inférieures ou égales à i et les classes ré-étiquetées comme *positives* seront les classes supérieures à i .

En suivant cette stratégie, nous obtenons 4 classifieurs binaires répartis comme ceci :

$$\begin{aligned}
 S_{\text{Avant contre Après}} &= \{Cl_1, Cl_2, Cl_3, Cl_4\} \\
 &= \{(\{C_1\}, \{C_2, C_3, C_4, C_5\}), \\
 &\quad (\{C_1, C_2\}, \{C_3, C_4, C_5\}), \\
 &\quad (\{C_1, C_2, C_3\}, \{C_4, C_5\}), \\
 &\quad (\{C_1, C_2, C_3, C_4\}, \{C_5\})\}
 \end{aligned}$$

Ils représentent l'ensemble de tous les classifieurs appris.

De cette manière, le classifieur Cl_i apprend à différencier les instances *négatives* appartenant aux classes allant de 1 à i face aux instances *positives* des classes de $i + 1$ à K . Donc, pour une instance x_j donnée, le classifieur Cl_i donnera la probabilité que cette instance appartienne à la classe *positive* (c'est à dire à l'ensemble des classes supérieures à i), notée $P(Cl_i|x_j)$.

Ainsi, la stratégie *Avant contre Après* nécessite peu de classifieurs et elle permet d'exploiter dans chacun des sous-classifieurs l'intégralité du jeu de données.

Après cette phase d'apprentissage, afin de prédire la classe d'appartenance d'une nouvelle instance x_j , les auteurs ont proposé une combinaison des probabilités des sous-classifieurs :

$$\begin{aligned}
 P(y_1|x_j) &= 1 - P(Cl_1|x_j) \\
 P(y_i|x_j) &= P(Cl_{i-1}|x_j) - P(Cl_i|x_j), 1 < i < K \\
 P(y_K|x_j) &= P(Cl_{K-1}|x_j)
 \end{aligned} \tag{4.7}$$

Cette première méthode de combinaison permet de donner des scores aux classes. La classe recevant la « probabilité » la plus élevée sera assignée à x_j .

Cependant, comme l'ont montré CARDOSO et COSTA (2007), cette approche peut produire des probabilités négatives. En effet, si pour la $i^{\text{ème}}$ classe, cette inégalité est vraie : $P(Cl_i|x_j) > P(Cl_{i-1}|x_j)$ alors la « probabilité » $P(y_i|x_j)$ sera négative. Ils ont conséquemment proposé une modification de la méthode de combinaison des probabilités des sous-classifieurs, permettant de résoudre ce problème :

$$\begin{aligned}
 P(y_1|x_j) &= 1 - P(Cl_1|x_j) \\
 P(y_i|x_j) &= (1 - P(Cl_i|x_j)) \prod_{n=1}^{i-1} (P(Cl_n|x_j)), 1 < i < K \\
 P(y_K|x_j) &= \prod_{n=1}^{K-1} (P(Cl_n|x_j))
 \end{aligned} \tag{4.8}$$

Les inconvénients de cette stratégie sont doubles. D'une part il y a les difficultés de compréhension de la fonction de coût global, comme indiqué par LI et LIN (2007) « Puisque les sous-problèmes sont résolus séparément et que la sortie est ensuite combinée, la généralisation de cette méthode est difficile à analyser ». D'autre part, l'équilibre entre les données traitées par les sous-classifieurs pose problème, comme l'ont mentionné CHANG, CHEN et HUNG (2011) « Dans cette approche, il n'y a pas de propriété sensible aux coûts ¹⁰ permettant de pondérer les classes en fonction de leurs tailles. ».

Nous allons voir dans la section suivante qu'il est important de prendre en compte l'existence de mesures d'évaluation adaptées aux problèmes ordinaux, c'est à dire de mesures prenant en compte la distance ¹¹ d'erreur de prédiction et/ou le respect de l'ordre ¹² des données réelles.

4.3 Méthodes d'évaluations de la classification ordinale

Par nature, les problèmes de classification ordinale ne doivent pas être évalués comme des problèmes de classification standard.

Dans les problèmes de classification standard, l'évaluation porte sur la quantité d'erreurs commises par le système sans se soucier du fait que les erreurs peuvent être plus ou moins proches (ou éloignées) de la valeur réelle. Ce qui est utilisé est strictement le nombre de cas correctement classés, quelle que soit l'importance des erreurs commises. Pourtant, un système de classification ordinale produisant des erreurs nombreuses mais légères semble préférable à un système produisant des erreurs de plus grande importance mais en moins grande quantité. L'objectif est de s'approcher de la valeur réelle tout en conservant une cohérence entre les prévisions faites. Les erreurs importantes et inconsistantes doivent être sanctionnées.

Nous pouvons résumer cela en relevant que dans la classification ordinale, il existe deux critères d'évaluation importants :

L'un est l'ordre absolu qui traduit l'exactitude et la proximité de la réponse. La méthode évaluée est-elle précise ? Les prédictions faites sont-elles proches des valeurs réelles ? Quelle est l'importance des erreurs commises ? Pour évaluer ce premier critère nous devons mesurer l'importance des erreurs, ce qui peut par exemple être fait en calculant la distance entre les valeurs prédites et les valeurs réelles.

10. En anglais : *Cost-sensitive property*

11. Comme le font les fonctions de coûts utilisées dans la régression *c.f.* sous-section 4.1.2.

12. Comme le font les fonctions de coûts utilisées dans le classement *c.f.* sous-section 4.1.3.

Le second critère d'importance dans les méthodes ordinales est l'ordre relatif. En effet, savoir si un système est précis est une bonne chose, mais il est également important de savoir si ce système préserve l'ordre pour un ensemble de cas prédits. L'algorithme développé est-il cohérent et préserve-t-il l'ordre pour un ensemble de cas prédits? Pour évaluer ce critère nous devons comparer deux listes ordonnées, l'une est la liste des données prédites, l'autre est la vérité terrain associée. Ces approches sont basées sur des comparaisons de distributions tels que les coefficients de corrélation.

Le reste de cette section traite des différentes mesures d'erreurs absolues (basées sur les instances et les distances) et relatives (basées sur le classement, l'ordre).

4.3.1 Mesures usuelles en classification

En classification standard une mesure d'évaluation très commune est la précision¹³ (*ACC*), qui peut être vue comme la proportion de documents bien classés pour tous les documents classifiés :

$$ACC(\hat{h}, T) = \frac{|\{x_i \in T : \hat{h}(x_i) = h(x_i)\}|}{|T|}, \quad (4.9)$$

où T représente l'ensemble de test. L'*ACC* est utilisée (KRAMER, WIDMER, PFAHRINGER et GROEVE, 2001 ; FRANK et HALL, 2001 ; POTHARST et BIOCH, 1999 ; WAEGEMAN et BOULLART, 2006) pour évaluer les problèmes ordinaux. Cette valeur reflète l'exactitude d'un système.

Avec un objectif inverse, « la version en miroir » de la précision est le taux d'erreur¹⁴ (DUI-VESTTEIJN et FEELDERS, 2008 ; SHASHUA et LEVIN, 2003), également appelé *Mean Zero-One Erreur (MZOE)* (CHU et GHARAMANI, 2004 ; CHENG, WANG et POLLASTRI, 2008) est la fraction de prévisions incorrectes (BACCIANELLA, ESULI et SEBASTIANI, 2009a) avec un coût égal pour toutes les erreurs :

$$MZOE(\hat{h}, T) = \frac{|\{x_i \in T : \hat{h}(x_i) \neq h(x_i)\}|}{|T|} = 1 - ACC(\hat{h}, T). \quad (4.10)$$

Cette mesure a encore d'autres appellations telles que : *Misclassification Error Rate (MER)* (COSTA, ALONSO et CARDOSO, 2008 ; COSTA, SOUSA et CARDOSO, 2010), *Zero-One Error* (CHENG, WANG et POLLASTRI, 2008), *classification error* (SHASHUA et LEVIN, 2003). Elle est aussi nommée R_{0-1} dans les travaux de HERBRICH, GRAEPEL, OBERMAYER et al. (1999). L'expressivité de cette valeur est parfois plus adaptée pour comparer les méthodes.

13. En anglais : *Accuracy*

14. En anglais : *Error rate*

Cependant, bien que largement utilisées, l'*ACC* et la *MER* ne sont pas en accord avec les préceptes des problèmes ordinaux. Ces valeurs ne reflètent pas bien ce que nous recherchons dans les problèmes ordinaux. En effet, lorsque l'on traite un problème ordinal, il est important de savoir si un système donné produit souvent des résultats proches des objectifs, plutôt que des résultats lointains et/ou irréguliers. La confiance dans l'approximation est ce qui motive l'utilisation de méthodes ordinales, et elle augmente si les éventuelles erreurs sont matérialisées par des résultats proches de la vérité terrain.

Dans le cadre de la classification ordinaire, la distance d'erreur est un paramètre essentiel, qui requiert des mesures d'évaluations adaptées.

4.3.2 Variantes tolérant certaines erreurs

Nous pouvons rechercher une interprétation plus souple des prévisions. GAUDETTE et JAPKOWICZ (2009) ont discuté de cet aspect et ont comparé d'autres mesures de classification ordinaire avec la précision à $1, 2 \dots n$ notée (ACC_n) qui est une variante de l'*ACC* qui tolère une plus large gamme de réponses valides, *i.e.*, une mauvaise classification de $\pm n$ classes est tolérée. Nous pouvons remarquer que la précision à 0 est équivalente à l'*ACC*.

$$ACC_n(\hat{h}, T, n) = \frac{|\{x_i \in T : |\hat{h}(x_i) - h(x_i)| \leq n\}|}{|T|} \quad (4.11)$$

Le Score Cumulé¹⁵ (*CS*) proposé par GENG, ZHOU et SMITH-MILES (2007) est inspiré par le concept de tolérance inclus dans l' ACC_n . Le *CS* est une version proportionnelle de l' ACC_n . Cette mesure est utilisée dans des travaux de CHANG, CHEN et HUNG (2011) pour évaluer des méthodes ordinales.

Le *CS* et l' ACC_n sont des mesures plus « grossières » que l'*ACC*. Leur tolérance témoigne déjà de l'intérêt porté à l'acceptation des erreurs se rapprochant des objectifs.

4.3.3 Mesures issues de la régression

À partir des fonctions de coût des méthodes de régression, les mesures de classification ordinaire ont hérité des mesures fondées sur la distance, telles que l'Erreur Absolue Moyenne (*MAE*¹⁶) et l'Erreur Quadratique Moyenne (*MSE*¹⁷). Ces mesures se sont popularisées au sein de la communauté de la classification ordinaire en raison de leurs propriétés intéressantes qui répondent aux objectifs de prise en compte de l'importance des erreurs, en plus de leur quantité.

15. En anglais : *Cumulative Score*

16. En anglais : *Mean Absolut Error*, également appelée : *Mean Absolute Deviation (MAD)*

17. En anglais : *Mean Square Error*

La *MAE* est une mesure d'erreur dans laquelle le coût associé aux erreurs augmente de manière graduelle en fonction de la distance entre la valeur prédite et la valeur réelle :

$$MAE(\hat{h}, T) = \frac{1}{|T|} \sum_{x_i \in T} |\hat{h}(x_i) - h(x_i)| \quad (4.12)$$

Cette mesure a été utilisée dans de nombreux travaux de classification ordinale (LI et LIN, 2007 ; WAEGEMAN et BOULLART, 2006 ; CHU et GHARAMANI, 2004 ; CHENG, WANG et POLLASTRI, 2008). Elle est parfois aussi utilisée et nommée Coût de Classement Moyen¹⁸ (SHASHUA et LEVIN, 2003 ; HARRINGTON, 2003 ; CRAMMER et SINGER, 2001). L'intérêt de cette mesure est qu'elle pénalise les erreurs de manière linéaire et graduelle. Plus les prédictions sont lointaines de la valeur réelle, plus l'erreur associée est importante.

La seconde mesure basée sur la distance qui provient de la régression est la *MSE* :

$$MSE(\hat{h}, T) = \frac{1}{|T|} \sum_{x_i \in T} (\hat{h}(x_i) - h(x_i))^2 \quad (4.13)$$

Cette mesure pénalise plus sévèrement les erreurs importantes en raison de l'élévation au carré des distances d'erreur (GAUDETTE et JAPKOWICZ, 2009). Cependant, cette mesure n'est pas autant utilisée que la *MAE*.

Ces méthodes amènent toutefois plusieurs questions. Est-ce que l'échelle linéaire ou quadratique a un sens avec les données de notre problème ? Est-ce que l'échelle choisie reflète bien le problème ? Avec ces mesures, les résultats dépendront des valeurs attribuées aux étiquettes des classes comme indiqué par CARDOSO et SOUSA (2011). Une possibilité est de choisir de convertir les étiquettes numériques en rang. De cette manière, potentiellement, les étiquettes attribuées ne sont pas les causes d'un biais dans le calcul de l'erreur (voir le tableau 4.1).

indices	1	2	3	4	5	6	...	n
scores	0.4	1.7	6.3	7.9	9.5	31.6	...	45.7
étiquettes numériques	1	1	8	8	8	34	...	42
étiquettes textuelles	Faible	Faible	Correct	Correct	Correct	Bon	...	Excellent
rangs	1	1	2	2	2	3	...	K

TABLEAU 4.1: Différences de conversion entre indices, scores, étiquettes et rangs (avec K classes et n instances).

Les résultats obtenus auront alors un sens en terme de distance de rang moyen. Cependant, du fait de l'élévation au carré, les résultats obtenus avec la *MSE* ne sont pas dans la même « dimension » que les rangs choisis, c'est pourquoi la *RMSE*¹⁹ est intéressante.

18. En anglais : *Average rank loss*

19. En anglais : *Root Mean Squared Error*

C'est en fait une variante de la MSE utilisant une racine carrée, $RMSE(\hat{h}, T) = \sqrt{MSE(\hat{h}, T)}$. Cette mesure est utilisée notamment dans les travaux de KRAMER, WIDMER, PFAHRINGER et GROEVE (2001) et nous pouvons remarquer qu'elle se comporte comme la MAE lorsqu'elle est utilisée avec des rangs. Elle donne ainsi des résultats interprétables en terme de variation moyenne de rangs (*i.e.*, à $\pm n$ rangs).

Pour convertir les étiquettes en rangs, une solution possible est d'utiliser une définition de ces mesures directement depuis la matrice de confusion (CARDOSO et SOUSA, 2011) :

$$MAE_{CM}(M) = \frac{1}{|M|} \sum_{r=1}^n \sum_{c=1}^n M_{r,c} |r - c|, \quad (4.14)$$

où M représente la matrice de confusion obtenue sur le jeu de données de test T . $M_{r,c}$ représente le nombre d'instances de la classe r classifiée dans la classe c . Dans cette mesure nous comparons les étiquettes transformées en rang ce qui permet d'obtenir un résultat interprétable en terme de $\pm n$ rangs.

4.3.4 Améliorations des mesures fondées sur la distance

Avec les données ordinales certaines considérations supplémentaires doivent être prises en compte. En effet, la nature des données ordinales pose problème car la plupart du temps, la population des classes des ensembles de données ordinales est très déséquilibrée. De ce fait, une mesure basée sur la distance comme la MAE ou la MSE va favoriser les classes les plus représentées. Le résultat se ressentira sur le classifieur ordinal appris qui sera un « classifieur trivial » minimisant l'erreur moyenne en votant principalement pour la classe la plus représentée. Pour répondre au problème soulevé par la taille des classes, BACCIANELLA, ESULI et SEBASTIANI (2009a) ont proposé d'utiliser les macro-moyennes de la MAE et de la MSE :

$$MAE^M(\hat{h}, T) = \frac{1}{n} \sum_{j=1}^n \frac{1}{|T_j|} \sum_{x_i \in T_j} |\hat{h}(x_i) - h(x_i)|, \quad (4.15)$$

où n est le nombre de classes et T_j l'ensemble de données de test appartenant à la classe y_j . De cette manière, ces mesures ne dépendent plus de la taille de la classe la plus représentée. Toutes les classes sont évaluées individuellement. Le résultat moyen ne souffre donc plus du déséquilibre entre les classes du jeu de données.

Dans le cadre de la classification ordinale de critiques de DVDs, GAUDETTE et JAPKOWICZ (2009) comparent un ensemble de méthodes d'évaluation couramment utilisées en classification ordinale (ACC , MSE , MAE , $Correl$, $NDPM$, ACC_n , $ACC + Correl$). La conclusion de leurs travaux indique que la MSE et la MAE sont les deux mesures les plus appropriées pour les problèmes ordinaux puisqu'elles sont capables d'évaluer les erreurs les plus faibles, comme le fait l' ACC à $1, 2 \dots n$, mais en une seule mesure. La mesure de performance de distance normalisée ($NDPM$ ²⁰) utilisée dans cette étude a été introduite par YAO (1995). La $NDPM$ est une mesure utilisée en recherche d'information pour évaluer la concordance et la discordance entre deux systèmes :

$$NDPM = \frac{1}{2}(1 + F - R). \quad (4.16)$$

F représente le taux de faux positifs²¹, soit le nombre de documents non pertinents retrouvés, divisé par le nombre total de documents non pertinents. F peut être vu comme la proportion de désaccord entre les deux systèmes comparés. R est le rappel, qui représente au contraire le nombre de documents pertinents trouvés divisé par le nombre total de documents pertinents. R peut être considéré comme la proportion d'accord entre les deux systèmes comparés. Les valeurs de la $NDPM$ vont de 0 à 1, et plus son score est bas, plus les systèmes sont en accord.

Dans cette étude comparative des mesures de classification ordinale (GAUDETTE et JAPKOWICZ, 2009), les auteurs observent que la $NDPM$ est adéquate avec la mesure de corrélation linéaire ($Correl$), que nous allons définir dans la sous-section suivante.

4.3.5 Mesures utilisant l'ordre relatif

Comme indiqué dans l'introduction du présent chapitre, il est intéressant de savoir si un système préserve l'ordre des prédictions. Nous voulons savoir si un système est consistant dans les prédictions qu'il réalise : est-ce que l'ordre local est toujours présent ? Pour évaluer l'ordre, les mesures d'évaluation basées sur le classement utilisent des paires d'instances. Nous voulons évaluer la relation entre deux distributions, la distribution des valeurs prédites et la distribution des valeurs réelles. Les mesures d'évaluation utilisant l'ordre relatif entre les instances sont conçues pour atteindre cet objectif. Elles sont basées sur des comparaisons de distributions tels que les coefficients de corrélation de rangs²². Une des manières les plus connues afin de comparer des distributions est la corrélation linéaire ($Correl$) :

$$Correl(X, Y) = \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_i (x_i - \bar{x})^2 \sum_i (y_i - \bar{y})^2}}, \quad (4.17)$$

20. En anglais : *Normalized Distance Performance Measure*

21. En anglais : *Fallout* parfois écrit *Fall-out*, également appelé *False Positive Rate (FPR)*

22. En anglais : *Rank correlation coefficients*

où \bar{x} et \bar{y} sont les moyennes respectives de x et y . Cette mesure est aussi appelée Corrélation de Spearman²³ (SPEARMAN, 1904), également définie comme le coefficient de corrélation de Pearson entre le rang des variables (MYERS, WELL et LORCH, 2010).

La mesure *Correl* est utilisée dans la classification ordinale (KRAMER, WIDMER, PFAHRINGER et GROEVE, 2001 ; COSTA, SOUSA et CARDOSO, 2010) et comparée à d'autres mesures ordinales (GAUDETTE et JAPKOWICZ, 2009 ; CARDOSO et SOUSA, 2011). La corrélation de Pearson a été utilisée comme fonction de covariance/de noyau dans l'approche de CHU et GHAMRAMANI (2004). Cependant, CARDOSO et SOUSA (2011) indiquent que cette mesure est toujours dépendante des valeurs de rang des étiquettes choisies pour représenter les classes.

Une mesure indépendante des valeurs choisies pour représenter les rangs des classes a été proposée afin d'évaluer la corrélation entre plusieurs distributions dans le cadre de la classification ordinale. Il s'agit du coefficient τ_b de Kendall (KENDALL, 1938), adapté et utilisé par COSTA, SOUSA et CARDOSO (2010). Ce coefficient évalue toutes les paires possibles d'instances $\binom{n}{2} = \frac{1}{2}n(n-1)$:

$$\tau_b = \frac{c - d}{\sqrt{c + d + e_t} \sqrt{c + d + e_p}}. \quad (4.18)$$

Pour toutes les paires d'instances, CARDOSO et SOUSA (2011) expliquent qu'ils observent la relation d'ordre (supérieure ou inférieure) au sein de la paire des classes réelles mais aussi des classes prédites. Si la relation d'ordre est identique, c'est une paire concordante qui compte pour c . Si elle est inversée c'est une paire discordante qui compte cette fois-ci pour d . Si les valeurs réelles des deux instances sont égales et que les valeurs prédites de ces deux mêmes instances sont égales (et non nécessairement conformes aux valeurs réelles), alors la paire n'est pas considérée. Sinon, si l'égalité n'est vérifiée qu'entre les valeurs réelles, c'est une paire véritable supplémentaire²⁴ qui compte pour e_t , et si à l'inverse l'égalité n'est vérifiée qu'entre les valeurs prédites, c'est une paire prédite supplémentaire²⁵ qui compte pour e_p .

En raison de la nature des problèmes ordinaux, il y a souvent peu de classes par rapport au nombre d'instances, le résultat se traduit par la présence de nombreuses valeurs égales qui sont présentes aussi bien dans l'ensemble des valeurs réelles associées aux instances que dans les valeurs prédites associées à ces mêmes instances.

23. En anglais : *Spearman's rank*

24. En anglais : *Extra true pair*

25. En anglais : *Extra predicted pair*

COSTA, ALONSO et CARDOSO (2008) ont proposé un coefficient pour comparer deux variables ordinales. Ce coefficient est basé sur les coefficients relationnels (LERMAN, 1992). Il fonctionne avec l'ensemble des paires qui vérifient la relation d'ordre « \leq » entre les valeurs de classes réelles S_1 et les valeurs de classes prédites S_2 :

$$r_{int} = A + B \frac{\text{card}(S_1 \cap S_2)}{\sqrt{\text{card}(S_1)\text{card}(S_2)}}. \quad (4.19)$$

Dans cette équation, A est fixé à -1 et B à 2 pour obtenir des valeurs de r_{int} comprises dans l'intervalle $[-1, 1]$. Cet indice de similarité ne fonctionne qu'avec les relations d'ordre. Toute l'information absolue est donc perdue.

Dans leur approche de la classification ordinale par paire, HERBRICH, GRAEPEL, OBERMAYER et al. (1999) évaluent leur système avec deux mesures ; la R_{0-1} que nous avons précédemment introduite (voir équation 4.10) et la R_{pref}^{0-1} , qui est une version de la R_{0-1} , calculée exclusivement sur les paires qui ont un rang différent. La R_{pref}^{0-1} peut donc être vue comme la proportion de paires mal classifiées.

La sensibilité $Sens_p^{p+1}$ et la spécificité $Spec_p^{p+1}$ (WAEGEMAN et BOULLART, 2006) sont deux mesures d'évaluation qui comparent les rangs de paires consécutives et donne un score en fonction de leur accord ou désaccord.

4.3.6 Mesure hybride

(CARDOSO et SOUSA, 2011) ont analysé et comparé certaines des mesures les plus utilisées en classification ordinale (r_{int} , MER , MSE , MAE , la corrélation de Spearman, le τ_b de Kendall). Ils proposent une nouvelle mesure, l'indice de classification ordinale²⁶, nommé OC_β^γ . Cette mesure (voir l'illustration 4.1) est conçue pour répondre aux différents inconvénients présents dans les mesures existantes.

$$OC_{\beta_1; \beta_2}^\gamma = \min \left\{ 1 - \frac{\sum_{(r,c) \in \text{path}} n_{r,c}}{N + (\sum_{\forall (r,c)} n_{r,c} |r - c|^\gamma)^{1/\gamma}} + \beta_1 \sum_{(r,c) \in \text{path}} n_{r,c} |r - c|^\gamma + \beta_2 N_{\text{disc_pos}} \right\}$$

ILLUSTRATION 4.1: Formule OC_β^γ . (source CARDOSO et SOUSA (2011))

26. En anglais : Ordinal classification index

CARDOSO et SOUSA (2011) critiquent en effet les mesures d'évaluation basées sur les instances²⁷ (*MER*, *MAE*, *MSE*...) car elles ne donnent pas d'information à propos de l'ordre relatif entre les classes prédites. Par opposition, ils critiquent aussi les mesures d'évaluation basées sur le classement²⁸ (corrélation de Spearman, τ_b de Kendall...), parce que les valeurs absolues des classes sont perdues. Ceci rend ces mesures plus adaptées aux problèmes de classement qu'aux problèmes ordinaux. C'est pourquoi ils ont conçu une nouvelle mesure qui satisfait conjointement à ces deux critères.

La mesure OC_{β}^{γ} peut être vue comme une solution hybride entre les mesures basées sur les instances (et les distances) et les mesures basées sur le classement. CARDOSO et SOUSA (2011) incorporent des éléments des mesures basées sur le classement dans une mesure basée sur les instances. La formulation complète de leur mesure ainsi qu'une explication étape par étape de la conception de leur mesure est disponible dans leur travaux (CARDOSO et SOUSA, 2011). Cette mesure fonctionne directement dans la matrice de confusion et peut être appliquée à une instance unique, contrairement aux mesures traditionnelles basées sur le classement.

Le paramètre γ pénalise la distance à la diagonale de la matrice de confusion. Le paramètre β contrôle le style de l'approche choisi. Plus le paramètre β est faible, plus l'approche s'apparente aux approches basées sur le classement.

4.3.7 Utilisation des méthodes d'évaluation en apprentissage supervisé

Il y a deux façons d'utiliser les mesures d'évaluation au sein des systèmes d'apprentissage supervisé.

La manière la plus simple consiste à les utiliser à l'extérieur du modèle, l'objectif étant de calibrer les paramètres du système grâce à elles. Cela peut être fait lors d'une étape de validation croisée. Les paramètres sont choisis parmi des plages de valeurs possibles et, à partir de toutes les combinaisons générées, les modèles sont évalués par la mesure sélectionnée. L'objectif est de trouver la meilleure combinaison de paramètres, c'est à dire celle qui donne les meilleurs résultats avec la mesure utilisée.

L'autre manière est d'utiliser la mesure à l'intérieur du modèle. Beaucoup de modèles utilisent une fonction de coût pour pénaliser les exemples. On peut choisir d'intégrer une fonction de coût particulièrement pertinente pour changer la fonction objectif final d'un nouveau système. Le système résultant sera alors plus adapté aux nouveaux objectifs ainsi intégrés dans le modèle.

27. En anglais : *Instance-based errors*

28. En anglais : *Ranking-based errors*

Bien sûr, ces mesures peuvent également être utilisées comme des critères externes d'évaluation de tous types de systèmes, et ainsi donner des scores d'évaluations comparatifs entre les systèmes.

4.4 Jeux de données ordinaux

Dans les problèmes ordinaux, nous souhaitons estimer une valeur pour laquelle un ordre peut être établi. Les données de ces problèmes doivent disposer d'un tel ordre. Plusieurs options ont été envisagées pour obtenir des jeux de données ordinaux, permettant l'évaluation des méthodes de classification ordinale.

Certaines approches de quantification ou de discrétisation partent d'ensembles de données de régression et convertissent la valeur numérique observée dans un ensemble ordonné de classes. Une possibilité est de choisir de créer des classes de même taille (*i.e.* , avec le même nombre d'éléments à l'intérieur de chacune des classes), ce qui produit des ensembles de données équilibrés. Cependant, si nous regardons les problèmes ordinaux du monde réel, cela n'est presque jamais le cas. Par exemple, si nous regardons un problème d'évaluation de films avec une notation de 1 à 5 étoiles, les évaluations moyennes de chaque film ne sont pas équilibrées : On ne trouve généralement pas le même nombre de films notés avec 1, 2, 3, 4 et 5 étoiles. En outre, il est nécessaire de choisir une « bonne taille » pour les classes car il est possible que l'information ordinaire des éléments ne soit pas significativement différente entre les classes d'une taille trop grande ou trop petite. Si nous disposons d'un *a priori* sur les données, nous pouvons sans doute estimer les meilleures coupes pour convertir notre problème en classes. Mais cela dépend de notre propre compréhension du comportement ordinal des données du problème particulier que nous voulons résoudre. De ce fait, notre jugement peut être fortement orienté par notre expérience personnelle. En revanche, si nous ne disposons pas d'*a priori*, nous pouvons choisir de faire des coupes régulières (*e.g.* , linéairement espacées) à l'intérieur de l'échelle ordinaire de notre problème. Les ensembles de données produits seront donc en quelque sorte moins prévisibles vis-à-vis de toute considération particulière que nous pouvons avoir sur le sujet.

Une autre approche possible est la génération de données de synthèse dans lesquelles nous disposons d'un contrôle total sur les données produites. Cette approche peut être utilisée dans le cadre du développement d'une méthode de classification ordinale afin de fournir une meilleure compréhension et une meilleure analyse du comportement de celle-ci.

Finalement, nous pouvons utiliser des données du monde réel naturellement ordonnées. Avec ce type de données, il n'est pas rare que l'ensemble de données soit impacté par un manque d'informations concernant certaines classes. Ce qui se traduit par des problèmes de déséquilibre entre les classes. La majeure partie de ces ensembles de données provient d'évaluations et de notes laissées par des utilisateurs. La plupart des systèmes d'évaluation proposent aux utilisateurs de donner une évaluation quantitative, exprimée sur une échelle ordinale. Ceci explique en partie l'engouement récent pour les méthodes de classification ordinale.

Nous présentons dans les sous-sections suivantes certains des jeux de données ordinaux les plus usités. D'autres jeux de données ordinaux sont répertoriés en annexe E à la page 177.

4.4.1 Jeux de données issus de la régression

Jeux de données de Weka

L'outil Weka (HALL, FRANK, HOLMES, PFAHRINGER, REUTEMANN et WITTEN, 2009) inclut 29 jeux de données de régression qui ont été collectés et mis à leur disposition par Luís Torgo²⁹ : *Abalone, Ailerons, Delta Ailerons, Elevators, Delta Elevators, 2D planes, Pole Telecomm, Friedman Artificial Domain, MV Artificial Domain, Kinematics of Robot Arm, Computer Activity (2), Census domains (2), Auto MPG, Auto Price, Boston housing, Diabetes, Pyrimidines, Triazines, Machine CPU, Servo, Wisconsin Breast Cancer, Pumadyn domains (2), Bank domains (2), California Housing, et Stocks domain.*

Un grand nombre d'entre eux provient du dépôt de *StatLib*³⁰ et du dépôt de *UCI ML*³¹. Plus d'informations peuvent être trouvées sur le site personnel de Luís Torgo³². Le téléchargement de ces jeux de données est aussi possible via le site de Weka dans la rubrique jeux de données³³.

Cet ensemble de jeux de données a été partiellement ou entièrement utilisé pour évaluer des méthodes de classification ordinale dans de nombreux travaux (WAEGEMAN et BOULLART, 2006 ; CHU et KEERTHI, 2005 ; FRANK et HALL, 2001 ; DUIVESTIJN et FEELDERS, 2008 ; LI et LIN, 2007 ; CHU et GHARAMANI, 2004 ; CHENG, WANG et POLLASTRI, 2008).

Comme indiqué dans l'introduction de cette section, une étape de quantification ou de discrétisation est nécessaire pour convertir les données d'un problème de régression en un problème ordinal, et les résultats dépendront en partie de ce processus de conversion.

29. <http://www.dcc.fc.up.pt/~ltorgo/> dernier accès août 2015

30. <http://lib.stat.cmu.edu/> dernier accès août 2015

31. <http://archive.ics.uci.edu/ml/> dernier accès août 2015

32. <http://www.dcc.fc.up.pt/~ltorgo/Regression/DataSets.html> dernier accès août 2015

33. <http://www.cs.waikato.ac.nz/ml/weka/datasets.html> dernier accès août 2015

CHU et GHAHRAMANI (2004) ont utilisé et converti ces données. Ils proposent leurs versions converties ainsi que les codes sources de leurs méthodes sur leur site web³⁴, afin de permettre la production de résultats comparatifs.

4.4.2 Jeux de données de synthèse générés

Un grand nombre d’auteurs (HERBRICH, GRAEPEL, OBERMAYER et al., 1999 ; POTHARST et BIOCH, 1999 ; CRAMMER et SINGER, 2001 ; HARRINGTON, 2003 ; CHU et GHAHRAMANI, 2004 ; WAEGEMAN et BOULLART, 2006 ; COSTA, ALONSO et CARDOSO, 2008 ; COSTA, SOUSA et CARDOSO, 2010 ; CARDOSO et SOUSA, 2011) a utilisé des jeux de données de synthèse, générés dans le but de se placer dans un cadre contrôlé au sein de la chaîne de traitement afin de permettre la validation *théorique* de leurs méthodes.

4.4.3 Jeux de données naturellement ordinaux

Jeux de données d’images

Il existe quelques jeux de données ordinaux d’images. Ces jeux peuvent concerner par exemple des problèmes d’estimation de l’âge des personnes, comme par exemple les jeux de données *MORPH Album 2*³⁵ et *FG-NET Aging Database*³⁶ qui sont utilisés dans les travaux de CHANG, CHEN et HUNG (2011). Dans leurs travaux, SHASHUA et LEVIN (2003) ont testé leur méthode de classification ordinale sur un jeu de données d’images représentant des photographies de voitures : « *Car size* ». Ce jeu de données d’images est composé de 3 classes ordinales représentatives des tailles des voitures et l’objectif de leur méthode était de déduire la taille d’une voiture à partir d’une photographie de celle-ci.

Jeux de données basés sur l’évaluation de biens et de services

Dans leur article, BLITZER, DREDZE et PEREIRA (2007) ont construit un jeu de données pour l’analyse adaptative aux domaines des sentiments. Leur jeu de données comporte une sélection d’évaluations de produits *Amazon*³⁷. Les évaluations sont échelonnées sur une échelle de 0 à 5 étoiles et concernent des livres, des DVDs, des appareils électroniques et des appareils de cuisine. Ce jeu de données est téléchargeable sur le site des auteurs³⁸.

34. <http://www.gatsby.ucl.ac.uk/~chuwei/ordinalregression.html> dernier accès août 2015

35. Payant via <http://www.faceaginggroup.com/morph/> dernier accès août 2015

36. http://fipa.cs.kit.edu/433_451.php, mais indisponible lors de nos tentatives depuis août 2015

37. <http://www.amazon.com> dernier accès août 2015

38. <http://www.cs.jhu.edu/~mdredze/datasets/sentiment/> dernier accès août 2015

Utilisé dans l'analyse des sentiments (polarité, degré et subjectivité), le jeu de données *movie-review*³⁹ a été utilisé par GAUDETTE et JAPKOWICZ (2009) pour la comparaison des méthodes d'évaluations utilisées en classification ordinale.

Le jeu de données *TripAdvisor 15763*⁴⁰ (*TripAdvisor Users review (15 763) dataset*) a été introduit par BACCIANELLA, ESULI et SEBASTIANI (2009b). C'est une collection d'avis en anglais laissés par des internautes concernant des voyages passés dans des hôtels des villes de Pise et de Rome. Les avis sont en texte brut et pour chacun d'entre eux, une note de 1 à 5 étoiles est associée. Cette base de données a été utilisée afin d'évaluer les jeux de données déséquilibrés par BACCIANELLA, ESULI et SEBASTIANI (2009a).

Un jeu de données de 5,6 millions d'avis laissés par des utilisateurs sur le site *Amazon* (*Amazon.com Users review (5.6 million) dataset*) a été collecté par JINDAL et LIU (2008). Ce jeu de données est utilisé pour la fouille d'opinions, l'analyse de sentiments, la recherche et la détection d'opinions parasites. Chaque avis est composé de 8 parties en texte brut incluant : la valeur de l'évaluation, l'identifiant du produit, l'identifiant de la personne ayant laissé l'avis, la date, le titre de l'avis, le corps de texte de l'avis, le nombre de remarques indiquant que l'avis a été utile et le nombre total de remarques. Ce jeu de données a été utilisé dans le même contexte que le jeu de données *TripAdvisor* par BACCIANELLA, ESULI et SEBASTIANI (2009a) pour évaluer le déséquilibre des jeux de données.

Nous mentionnons ici un dernier jeu de données de ce type, *Each Movie*, qui est le résultat des avis collectés par le système de recommandations du groupe de recherche *GroupLens Research* de *HP/Compaq Research* (anciennement, *DEC*). Le jeu de données produit contient 2 811 983 évaluations de films, portant sur 1 628 films, et fournis par 72 916 utilisateurs. Ce jeu de données a été utilisé dans de nombreuses expériences de classification ordinale (SHASHUA et LEVIN, 2003 ; HARRINGTON, 2003 ; CRAMMER et SINGER, 2001 ; CHU et GHARAMANI, 2004). Il a malheureusement été retiré par HP en octobre 2004⁴¹.

Ce jeu de données a été utilisé de nombreuses fois avec des méthodes de classification ordinale (SHASHUA et LEVIN, 2003 ; HARRINGTON, 2003 ; CRAMMER et SINGER, 2001 ; CHU et GHARAMANI, 2004).

39. <http://www.cs.cornell.edu/people/pabo/movie-review-data/> dernier accès août 2015

40. <http://tripadvisor.com> dernier accès août 2015

41. « *HP/Compaq Research (formerly DEC Research) ran the EachMovie movie recommender. When EachMovie was shutdown, the dataset was available to the public for use in research. MovieLens was originally based on this dataset. It contains 2,811,983 ratings entered by 72,916 for 1628 different movies, and it has been used in numerous CF publications. As of October, 2004, HP retired the EachMovie dataset. It is no longer available for download.* » — GroupLens Research.

4.5 Discussion

« Pourquoi la classification ordinale est elle plus difficile que la régression ? » Bien que tous les problèmes ordinaux partagent la même propriété d'ordre entre les classes, ils ne peuvent pas être interprétés de la même manière. En effet, la justification de la relation d'ordre est particulière et spécifique à chacun des problèmes, cela dépendra de la nature de l'ordre représenté.

Notre hypothèse est que les photographies contiennent des informations temporelles qui au fil des années ont évolué : les couleurs présentes, le point de vue, le style de prise de vue, le traitement photographique... En outre nous pensons que cette information temporelle peut être utilisée pour prévoir les dates de prises de vue des photographies.

Nous proposons de répondre à ce problème par l'utilisation de méthodes de classification ordinale dans le cadre de la classification visuelle de photographie dans le chapitre 6. En effet, la nature ordinale du temps doit être prise en compte pour améliorer les performances des techniques répondant à ce type de problème. Voilà pourquoi nous pensons que les méthodes ordinales sont plus adaptées et prouveront leur utilité dans nos expériences.

Nous nous situons au niveau des approches de la classification ordinale qui décomposent le problème ordinal en sous-problèmes binaires que nous avons présentées dans la sous-section 4.2.3. De part leur ressemblance aux méthodes de ré-étiquetages couramment utilisées dans la classification multi-classes (*Un contre Un* et *Un contre le Reste*), nos préférences se sont orientées vers cette stratégie ordinale du fait de sa logique... Nous intervenons sur les agencements de ces sous-problèmes au moment d'obtenir la prédiction car selon nos observations, les combinaisons que nous avons présentées dans les formules 4.7 et 4.8 peuvent être améliorées, et c'est ce que nous proposons dans le chapitre 6.

Chapitre 5

Constitution d'un jeu de données de photographies datées et contextualisées

Sommaire

5.1 Recherche de bases de photographies datées et contextualisées . . .	72
5.2 Constitution du corpus web-photo 1900-2000	81
5.3 Conclusion	86

Nous venons de voir de nombreux jeux de données utilisables pour évaluer les approches de classification ordinale. Il leur manque cependant à tous un ingrédient essentiel : des photographies datées¹. Dans ce chapitre nous présentons des jeux de données d'images datées et détaillons les raisons qui nous ont conduits à constituer à un corpus d'étude pour permettre l'évaluation de la datation de photographies grâce à leur contenu textuel et visuel. Ce corpus est décrit en détail en fin de chapitre.

Contexte et objectif

Le jeu de données de PALERMO, HAYS et EFROS (2012), déjà décrit en section 3.2, contient des images couleurs sur une période de 50 ans. Il permet d'expérimenter grâce au contenu visuel, la datation de photographies réparties équitablement selon 5 décennies contiguës. Sur ce sujet, nous avons conçu une amélioration des modèles de datation de photographies, qui est l'objet du chapitre 6.

1. Certains jeux contiennent des images de voiture et de personnes mais pas de photographies datées.

Nous avons mis en évidence dans les chapitres 2 et 3 le fort lien existant entre les photographies et le texte les environnant. Par conséquent, étant donné le rôle clé que joue le contenu textuel, il serait pertinent et judicieux d'entrevoir la datation de photographies grâce à celui-ci. Au regard de nos objectifs, le jeu de données de PALERMO, HAYS et EFROS (2012) n'offre pas cette possibilité.

Par ailleurs, rappelons qu'il existe des jeux de données purement textuels basés sur la datation des écrits, qui ont notamment été utilisés dans les travaux de JATOWT, AU YEUNG et TANAKA (2013) pour retrouver la période temporelle concernée par un écrit². C'est pourquoi, nous avons décidé de proposer à la communauté de la recherche d'information et de la vision par ordinateur, un jeu de données offrant simultanément ces deux aspects. Celui-ci permet pour la première fois d'évaluer des méthodes de datation de photographies utilisant le contenu textuel environnant.

5.1 Recherche de bases de photographies datées et contextualisées

Dès nos premières recherches, cette tâche s'est avérée difficile. Tout d'abord, trouver une base d'images cohérentes et exploitables est déjà problématique. Ensuite, si cette base doit contenir des photographies datées dans de bonnes proportions, à des fins d'analyse, cela devient plus compliqué. Enfin, une telle base doit en outre incorporer des informations textuelles décrivant ces images ou du moins ayant un lien avec celles-ci.

Nous avons choisi de considérer le spectre temporel le plus large possible pour une grande quantité d'images pour lesquelles : les dates de prises de vues étaient connues, la précision temporelle était de l'ordre de l'année et un texte environnant était disponible. Nous avons aussi pris en compte la distribution temporelle des sujets observés, afin d'éviter tout biais lié au contenu des images.

2. En anglais : *Document focus time*

5.1.1 Bases d'images datées

Nous avons trouvé un certain nombre de bases de données d'images datées disponibles via le web. Nous donnons à présent un ensemble de bases d'images que nous avons sélectionnées. Certaines de ces bases d'images sont volumineuses et concernent de multiples sources telles que des peintures, des photographies, des sculptures, *etc.* Elles couvrent des sujets très variés et contiennent des dates de prise de vue et parfois du texte. Pour certaines, un biais lié aux sujets est présent et pose problème car les modèles appris sur ces données risquent d'en être affectés. Dans de telles circonstances, il est systématiquement nécessaire de filtrer, trier, organiser et annoter la collection afin de pouvoir l'exploiter par la suite dans de bonnes conditions.

National Portrait Gallery

La base *National Portrait Gallery* est une base de portraits sur supports variés. Nous y trouvons des peintures, sculptures, dessins, miniatures, silhouettes, croquis, daguerréotypes, et autres, qui sont bien souvent datés. Malheureusement les textes associés aux images de cette base sont principalement des légendes donnant des indications techniques du support. Le problème de cette base réside aussi dans la trop grande diversité des supports. Avec ce genre de base, sans filtrage des données, il sera impossible de calibrer une méthode visant à dater uniquement des photographies. De plus, le risque d'utiliser une base de portraits tend plutôt à « l'apprentissage » de l'âge des sujets qu'à celui des informations temporelles des photographies (CHANG, CHEN et HUNG, 2011). Supposons par exemple que les photographies prises avant 1900 soient principalement des photographies de vieilles personnes et qu'au contraire celles prises après 1900 soient des photographies de jeunes personnes. Alors, si nous voulons évaluer une nouvelle image avec des méthodes calibrées sur cette base, l'attribution de la date sera largement influencée par la personne représentée et donc par son âge, au lieu d'être basée sur le contenu temporel décrit par l'aspect ancien ou moderne de la photographie. Une impression d'écran présentant l'interface web de « National Portrait Gallery » est disponible en annexe dans l'illustration F.1 page 181.

Bibliothèque de Toulouse

La base de la *Bibliothèque de Toulouse – Photographies anciennes* – contient 4 690 photographies datant des années 1857 à 1940. Les sujets sont variés : scènes de rue, cartes postales, quartiers de Toulouse, et les informations concernant le support sont relativement complètes, détails, dimensions de l'épreuve, etc. Malheureusement pour notre étude, cette base ne contient pas assez d'images et ne couvre pas une période temporelle suffisamment représentative, car les derniers clichés datent de 1940. Par ailleurs, il n'y a pas de texte associé à ces photographies. Ces limitations étant bloquantes, nous ne détaillerons pas davantage cette collection. Une photographie numérisée de la base de la bibliothèque de Toulouse et les détails techniques liés à celle-ci sont visibles en annexe, dans les illustrations F.2 et F.3 page 182.

Bibliothèque numérique Gallica

« Gallica est la bibliothèque numérique de la Bibliothèque nationale de France et de ses partenaires. En ligne depuis 1997, elle s'enrichit chaque semaine de milliers de nouveautés et offre aujourd'hui accès à plus de 2 millions de documents »³, (*Bibliothèque numérique Gallica*).

Cette base contient énormément de documents. Nous y trouvons notamment plus de 500 000 images en très grande partie datées. Celles-ci représentent essentiellement des gravures, peintures, dessins, photographies, carnets de photographies, illustrations de presse, etc. Le volume conséquent de cette base est un atout. Cependant la variété des supports présents et la non disponibilité des données textuelles contextualisantes représente une faiblesse au regard de nos objectifs. Des impressions d'écran présentant la base Gallica sont disponibles en annexe dans les illustrations F.4 et F.5 page 183.

Médiathèque de l'Architecture et du Patrimoine

« La Médiathèque de l'architecture et du patrimoine (MAP) a pour mission de collecter et constituer, classer et conserver, étudier, communiquer et valoriser :

- les archives et la documentation de l'administration des Monuments historiques ;
- le patrimoine photographique de l'État, riche d'au moins quatre millions de tirages et quinze millions de négatifs remontant aux origines de la photographie »⁴, (*Médiathèque de l'Architecture et du Patrimoine – Base Mémoire : Archives photographiques*).

3. Source : <http://gallica.bnf.fr/> dernier accès août 2015

4. Source : <http://www.mediathèque-patrimoine.culture.gouv.fr/> dernier accès août 2015

Elle propose plus de 600 000 images accessibles en ligne. Cette base vise à rassembler un maximum d'informations sur le patrimoine et tout ce qui s'y rattache : églises, chefs d'œuvre architecturaux, bâtiments, châteaux... Par exemple, pour certaines photographies, il peut s'agir de détails architecturaux d'un bâtiment précis (vitrail, arche, etc.) ou bien encore des photographies de paysages... Ce type de photographies est cependant difficile à dater, même pour un humain. De plus, dans cette base, les dates sont parfois absentes ou encore stockées sans être normalisées. Nous trouvons par exemple ces types de notations : « 1896.09.09 », « 1928-1929 », « 1900 ou 1910 », « 1963 », « 1965 (été) », « 1967.05 », « 1964 (pâques) », « 1928 08 01 04 et 07 », « 1929 08 30 et 31 et 1929 09 03 », « 1949 ? », « 1977-1978-1979 ? », « 1925.04.201925 », « 1890 (août à novembre) », « 1930 (années) », « 1981 (antérieur à) », « après restauration », « 1976 (Cliché antérieur à) », « Déblaiement », « Sans date », « 1929.05 (supposé) », « » et bien d'autres.

Il est par conséquent difficilement concevable d'utiliser ces informations temporelles sans un travail de nettoyage et de normalisation préalable, qui n'est pas exempt de difficultés ni de risques. L'illustration F.6 en annexe à la page 185 présente une impression d'écran de la base de la MAP.

Collection flickr « The Commons »

Le site flickr propose de partager des images. Il contient des photographies d'utilisateurs amateurs et professionnels mais aussi des collections de fonds numérisés en provenance d'archives nationales et de musées à travers le monde dans le cadre de la collection flickr « The Commons ».

La collection flickr « The Commons »⁵ a été lancée le 16 janvier 2008, avec pour objectif principal de favoriser l'accès aux classeurs détenus par des institutions, et de permettre au public d'apporter des informations complémentaires.

Un extrait de photographies appartenant à cette base est visible en annexe dans l'illustration F.7 à la page 186. Quelques institutions parmi la centaine de participantes au projet « The Commons » sont représentées dans l'illustration 5.1.

Les institutions participantes mettent des collections à disposition des utilisateurs, afin qu'elles soient commentées dans le but de les redocumentariser. Cette collection nommée « The Commons » est volumineuse et bien souvent les clichés anciens de cette base disposent de dates de prise de vue. En outre, cette collection est très peu bruitée en terme de photographies incongrues, d'images doublons, de faux, etc. Ceci provient du fait qu'elle est constamment alimentée par des institutions et non par des particuliers.

5. Source www.flickr.com/commons/ dernier accès août 2015

Institutions participantes

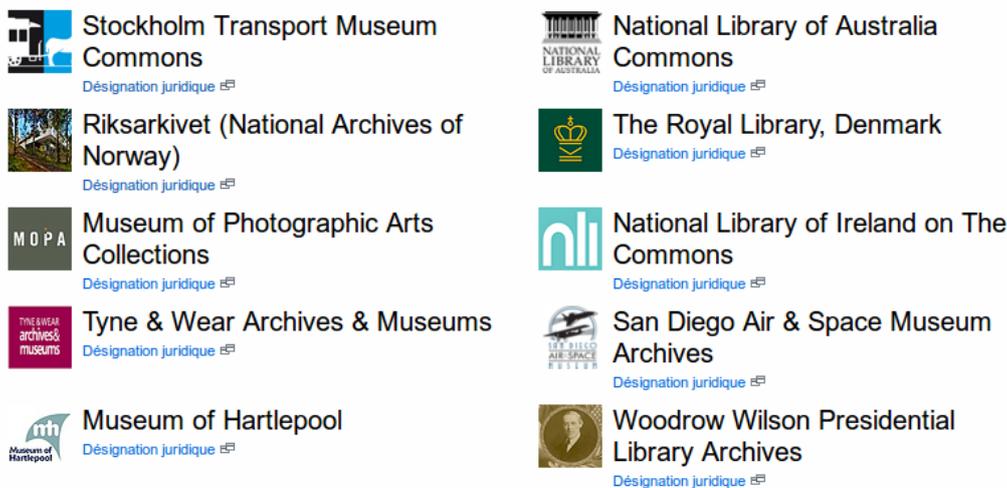


ILLUSTRATION 5.1: Quelques institutions participant au projet « The Commons ».

Le but de « *The Commons* » est d'augmenter l'accessibilité au grand public, via le web et les réseaux sociaux, de collections qui n'étaient jusqu'alors pas disponibles de cette manière. Cette collection favorise et compte sur la myriadisation⁶ aussi appelée redocumentarisation car c'est le public qui apporte l'information sur les clichés lorsqu'il le souhaite. Elle contient plusieurs millions de photographies datées remontant aux origines de la photographie dans les années 1820 et ce chiffre ne cesse de croître. Les thèmes de cette collection sont variés : portraits, paysages, monuments, faits historiques, objets *etc.* Les dates associées aux photographies répondent à des normes, ce qui est un net avantage face aux autres bases citées précédemment. De plus, les images de la collection « *The Commons* » disposent de données textuelles.

D'ailleurs, *flickr* met à la disposition du public une *API*⁷ qui permet d'accéder aux images et aux informations associées à ces dernières grâce à un langage de programmation, de manière simplifiée.

Remarquons que cette ressource a été exploitée par PALERMO, HAYS et EFROS (2012) dans la réalisation de leur jeu de données. Ils ont sélectionné 1 325 images couleurs parmi les photographies de la collection « *The Commons* ». Leur sous-collection permet d'expérimenter la datation de photographies par le biais des caractéristiques visuelles. Cependant, les données textuelles telles que les titres, les *tags* et d'autres informations, bien qu'accessibles, n'ont pas été exploitées.

6. Myriadisation (SAGOT, FORT, ADDA, MARIANI et LANG, 2011), en anglais : *Crowdsourcing*

7. *Application Programming Interface* : interface de programmation

Photographies issues de différentes sources du web

Le web est un terrain d'accueil pour tous les contenus, y compris les photographies datées et contextualisées. En effet, le web peut répondre aux besoins de chacun en matière d'archivage de contenus. Les normes et les standards existent mais ils sont multiples et évoluent sans cesse. Que l'on se place du côté de l'internaute créateur ou consommateur, le web est varié et plus ou moins structuré. C'est une expression de sa popularité et de son expansion non contrôlée. Utiliser le web comme matière première est potentiellement favorable à nos objectifs, car nous avons la possibilité de recueillir les données de nombreux sites web proposant des photographies. Certains de ces clichés seront ainsi suffisamment documentés pour offrir des contenus visuels et textuels. Compte tenu de la structuration des données web, une telle approche nécessite néanmoins l'usage de moyens de filtrage et de nettoyage des données. En effet, parfois les pages web ne sont pas valides (balises incomplètes, mauvais encodage des caractères, usage de technologies dépassées. . .) mais avant tout chaque site web est, toutes proportions gardées, singulier tant dans sa forme que dans son contenu. Ce qui entraînera par exemple des pertes d'informations lors de la manipulation de celles-ci (phases de détournage (LEJEUNE, BRIXTTEL et LECLUZE, 2015 ; LEJEUNE, 2013) via des expressions régulières ou d'extraction d'information en suivant les chemins *xpath*, etc.). De plus, une phase d'annotation de photographies sera inévitable pour obtenir, valider et normaliser les dates de prises de vue des clichés. Nous présentons en annexe l'exemple d'un site web contenant une image ancienne ainsi que le texte de l'article associé (illustration F.8 page 187). Ainsi le web peut répondre à nos exigences en matière de photographies datées sur un large spectre temporel mais aussi au regard des données textuelles. Ces avantages viennent au prix d'annotations qui doivent être effectuées manuellement et ajoutées.

5.1.2 Expérimentations à la réalisation d'un premier jeu de données

Parmi ces ressources d'images, certaines ne couvrent qu'une faible période de temps et ne disposent que de très peu de clichés. Les supports à l'origine de l'image numérisée sont parfois trop variés : peintures, sculptures, photographies, détails architecturaux etc. D'autres ne sont pas assez diversifiés en termes de sujet ou de thème photographique (contenant par exemple uniquement des portraits). En outre, les données textuelles sont quasiment inexploitable dans la plupart des cas.

Néanmoins, la collection *flickr* « *The Commons* » est une candidate intéressante pour notre sujet de recherche puisqu'elle dispose de nombreuses caractéristiques avantageuses. Ceci nous a incité à exploiter les données textuelles et visuelles provenant de cette collection. En comparaison des caractéristiques du jeu de données de PALERMO, HAYS et EFROS (2012), notre objectif est de récolter des photographies datées mais aussi les données textuelles associées, et ce, sur un spectre temporel plus large.

Nous avons par conséquent téléchargé⁸ un ensemble de documents numériques dans lesquels :

- les images pour lesquelles les dates étaient présentes, et situées entre 1820 et 1999⁹, sont stockées dans leur plus grande résolution en *jpg*¹⁰ ;
- les informations textuelles associées à ces images sont collectées et structurées au sein d'un fichier *XML*.

Cette entreprise nous a permis d'obtenir 42 310 documents auxquels sont associés image et texte. Les données que nous avons collectées en complément du contenu visuel pur sont les suivantes :

owner : l'identifiant unique de l'utilisateur ayant mis les images en ligne ;

id : l'identifiant unique de l'image sur le site *flickr*. Il nous permet de conserver le lien entre l'image et le fichier « *XML* » que nous avons créé, mais aussi d'avoir une référence vers les données liées à cette image sur le site *flickr*. Grâce à cet « *id* » nous pouvons aller rechercher lorsque nous le désirons, de nouvelles données concernant par exemple les commentaires des utilisateurs qui eux, évoluent perpétuellement ;

title : le titre de la photographie ;

description : la description de la photographie peut contenir beaucoup d'informations intéressantes mais la qualité et la quantité varient selon les utilisateurs. Certains fournissent le nom du photographe, ainsi que bon nombre d'informations concernant les circonstances de la prise de vue ;

dateupload : la date à laquelle la photographie a été mise en ligne, qui est indiquée selon un horodatage UNIX, c'est-à-dire un entier non signé correspondant au nombre de secondes écoulées depuis le 1^{er} janvier 1970 GMT ;

datetaken : dans *flickr*, il y a donc 2 dates différentes, la date de mise en ligne sur le site « *dateupload* » et la date de prise de vue de la photographie « *datetaken* ». C'est cette dernière qui nous intéresse.

La date de prise de vue est par défaut extraite des données EXIF¹¹ de l'image transmise à *flickr*, mais elle peut être ajustée par l'utilisateur.

Dans le cas d'images numérisées, la date de prise de vue par défaut sera donc celle de la numérisation par scanner ou par l'appareil photographique numérique ayant effectué l'acquisition. Cette date de prise de vue doit alors être changée manuellement car elle ne correspond pas à la date de prise de vue réelle.

8. À l'aide d'un *wrapper* python utilisant l'API *flickr* (MIGNON, 2015).

9. De nombreuses photographies ont été prises après 1999, chiffre qui est dû aux images pour lesquelles les dates n'ont pas été fixées et où la date de prise de vue est celle de la date de numérisation. Beaucoup d'appareils lorsqu'ils sont remis à zéro (piles changées, coupure de courant, etc.), disposent d'une date par défaut qui est le 1^{er} janvier 2000. C'est pourquoi, sur les sites de partage de photographies, beaucoup d'images ont prétendument été prises le 1^{er} janvier 2000. Afin d'éviter tout bruit dû à ce détail technique, nous avons choisi de nous limiter aux clichés dont la date est antérieure à l'an 2000.

10. *jpg* est un format de stockage d'images numériques

11. Les données EXIF sont des métadonnées associées aux images, qui sont enregistrées au moment de son acquisition numérique.

`datetakengranularity` : la date « `datetaken` » dispose aussi d'un paramètre de granularité¹² nommé « `datetakengranularity` ». Ce paramètre définit la précision dont nous disposons quant à la date de prise de vue originale. Ainsi pour une photographie prise le 7 avril 2012 à 09 :52 :17 selon la précision de l'information nous pouvons avoir des granularités différentes¹³ :

– granularité 0 : précise à la seconde près

e.g. : `taken="2012-04-07 09:52:17" takengranularity="0"`

– granularité 4 : durant le mois mentionné

e.g. : `taken="2012-04-01 00:00:00" takengranularity="4"`

– granularité 6 : au cours de l'année mentionnée

e.g. : `taken="2012-01-01 00:00:00" takengranularity="6"`

– granularité 8 : autour de l'année spécifiée (\pm quelques années)

e.g. : `taken="2012-00-01 00:00:00" takengranularity="8"`

`ownername` : le nom de l'utilisateur sur le site de *flickr* ;

`originalformat` : le format original de l'image qui a été envoyé sur *flickr* par l'utilisateur : *Raw, jpg, png, etc.* ;

`latitude` : coordonnée GPS ;

`longitude` : coordonnée GPS ;

`accuracy` : information quant à la précision des coordonnées GPS ; plus le chiffre est grand plus l'information est précise. La précision mondiale est 1, celle d'un pays est environ de 3, celle d'une région 6, d'une ville 11, et d'une rue environ 16. Les valeurs s'échelonnent de 1 à 16. La valeur par défaut est 16 ;

`tags` : les *tags* des mots-clés qui servent à décrire la photographie. Cette appellation regroupe les « *tags* » et « *machine tags* ». Ces derniers sont des *tags* organisés pour décrire plus finement le contenu plutôt que d'utiliser des termes vastes et génériques. Ils peuvent être utilisés au sein de groupes *flickr* pour définir des normes de nommage des *tags* et permettent ainsi une recherche plus aisée du contenu partageant des caractéristiques communes ;

`o_width` : largeur de l'image originale ;

`o_height` : hauteur de l'image originale ;

`pathalias` : nom de l'utilisateur utilisé dans l'*url*.

Un document numérique extrait de cette base est consultable en annexe page 189. Ce document est composé d'une image en *jpg* et d'un fichier descriptif en *XML*.

12. Les détails concernant ce paramètre sont accessibles à l'adresse suivante : <http://www.flickr.com/services/api/misc.dates.html> dernier accès août 2015

13. Nous pouvons remarquer que la granularité 2 (précise au jour près) n'est pas utilisée.

5.1. Recherche de bases de photographies datées et contextualisées

Dans les illustrations G.4, G.5 et G.6 page 192, nous représentons en annexe, par une série d'histogrammes, les caractéristiques générales de la répartition des données que nous avons récoltées. Comme nous pouvons l'observer, la répartition des images par année est fortement biaisée dans notre base. Elle contient tout de même 42 310 documents numériques composés de données textes et d'images datées, dont 8 831 disposent d'une date de prise de vue avec une « *datetakengranularity* » égale à 6 au maximum, c'est à dire précise au moins à l'année. Afin de filtrer le type de photographies présentes dans la base nous avons conçu un système d'annotation (c.f. l'illustration G.3 page 191 en annexe) pour extraire un sous-ensemble de 1 093 clichés. Le but d'une telle sélection était d'obtenir des photographies « datables », c'est à dire, des photographies ayant un contenu temporel. De même que JATOWT, AU YEUNG et TANAKA (2013) différencient les contenus temporels des contenus atemporels des textes, nous opposons les contenus visuels de la même manière. Pour ce faire, nous avons décidé de sélectionner les photographies de rues dans lesquelles une activité humaine était visible car des détails et des caractéristiques temporelles y sont présentes. Ces détails stéréo-typiques correspondent par exemple à l'architecture, aux voitures, aux vêtements, aux affichages, aux infrastructures et autres... À la différence des photographies de portraits ou de paysage pour lesquelles le contenu photographique n'est pas forcément temporel, nous favorisons dans notre étude les contenus humainement datables.

Voici un exemple d'images de rues que nous avons extraites, présentées sous forme de frise chronologique (voir l'illustration 5.2).



ILLUSTRATION 5.2: Une frise chronologique d'images de rues de notre base.

À la différence du jeu de PALERMO, HAYS et EFROS (2012), notre base contient des images en couleur, noir et blanc, sépia dont les photographies datent des années 1820 à 1999. Après 1999, la majorité des photographies numérisées proposées par les musées sont datées par la date de leur numérisation et non par la date réelle de leurs prises de vues. Nous les avons donc ignorées. Étant donné la richesse et l'étendue de cette base, elle constitue une contribution importante pour la datation d'images.

Suite à l'analyse de cette base dans son ensemble, nos observations ont montré que les données textuelles de celle-ci n'apportaient que peu d'informations contextualisantes car il s'agit dans la plupart des cas de légendes rédigées par les musées pour décrire le média et son support original. La légende existe mais son texte est peu exploitable car c'est une donnée technique qui ne s'inscrit pas en temps que donnée contextualisante portant une information temporelle.

Nous pouvons séparer ces deux cas d'usage de la photographie en contexte comme nous les avons indiqués dans nos hypothèses (voir sous-section 1.2.1). L'un étant l'usage du texte comme description technique de la photographie et l'autre étant la photographie comme l'illustration d'un propos textuel. C'est au fil de l'étude de ces données que cette distinction est apparue, nous faisant comprendre qu'il serait pertinent d'extraire et d'utiliser les documents se trouvant dans le second cas.

Du fait de cette distinction nous avons décidé de ne pas exploiter davantage cette ressource et d'en créer une nouvelle.

Il est à noter que cette collection reste toutefois exploitable dans un contexte de datation d'images visuelle sur un plus large spectre temporel que celui proposé par PALERMO, HAYS et EFROS (2012). Notons aussi que les classes que nous avons proposées sont des années, ce qui permet de faire varier la granularité de l'information temporelle (années, décennies ou autre (DIAS, MORENO, JATOWT et CAMPOS, 2012)). La distribution des images dans la collection « *The Commons* » est néanmoins fortement biaisée sur certaines années. Ceci est en partie dû aux collections historiques portant sur des faits marquants, qui ont lieu sur des périodes de temps restreintes, et durant lesquelles de grande quantité de photographies ont été accumulées par les musées. Ceci induit que le jeu de données est composé de classes non balancées.

Le web, quant à lui, offre une variété et une absence de normalisation des dates qui peuvent compliquer l'extraction de telles données hétérogènes. En revanche, le web offre des données textuelles riches qui constituent de solides contextualisations de photographies associées. C'est pourquoi nous avons orienté notre recherche vers cette ressource, en dépit du fait qu'il nous sera nécessaire de sélectionner, filtrer et d'annoter ces contenus pour les expérimentations futures.

5.2 Constitution du corpus web-photo 1900-2000

Faisant suite aux observations reportées dans la section précédente, nous avons choisi d'utiliser le web, avec ses avantages et inconvénients, comme matière première pour la conception d'un jeu de données d'évaluation de la datation temporelle de photographies qui soit capable de prendre en compte à la fois leurs caractéristiques visuelles et leurs environnements textuels.

5.2.1 Collecte d'informations issues du web

Nous nous sommes concentrés sur la collecte de documents multimédia regroupant texte et image associés, traitant de villes et ayant un aspect temporel. Nous avons donc collecté un ensemble de pages web, récupérées grâce à *Google images* en utilisant des requêtes temporelles. Pour construire ces requêtes, nous avons sélectionné 50 villes¹⁴, correspondant aux 10 destinations les plus prisées de 5 régions du globe selon une étude menée par *MasterCard*¹⁵. Nous avons donc recherché des pages web contenant des images de ces villes en suivant l'idée de DIAS, MORENO, JATOWT et CAMPOS (2012) pour construire les requêtes temporelles. L'extension de requêtes peut-être une façon d'intégrer de l'information connexe à notre recherche (LOUPY, BELLOT, EL-BÈZE et MARTEAU, 1998 ; XU et CROFT, 1996). L'aspect temporel fait partie de ces besoins d'informations qui peuvent être ajoutés à chaque terme recherché. Nous avons choisi pour la constitution de notre jeu de données d'ajouter une année dans une plage de temps allant de 1820 à 2010. Dans notre étude, les termes étaient représentés par les 50 villes citées précédemment. Les requêtes sont donc ainsi définies :

une requête temporelle = « un nom de ville »_U« une année »

Nous avons utilisé l'ensemble de ces requêtes temporelles et collecté les informations concernant les résultats de ces recherches :

- les images ;
- les pages web dont proviennent ces images ;
- le classement ;
- le *snippet* retourné par *Google images* contenant l'adresse de l'image, l'adresse de la page originale, etc.

Nous avons ainsi collecté un total d'environ 150 000 documents numériques. Nous présentons un exemple d'un document ainsi que les données qui lui sont associées dans l'annexe H.2 à la page 197. Ceci constitue un corpus volumineux qu'il est nécessaire de filtrer pour obtenir des résultats de bonne qualité. Comme tout résultat de moteurs de recherche, les sujets retrouvés par nos requêtes temporelles ne sont pas toujours pertinents. Le corpus filtré devra ensuite être annoté afin de le rendre exploitable dans le contexte de nos travaux.

14. La liste complète est indiquée en annexe dans le tableau H.1 page 196.

15. Source *MasterCard* : l'étude repose sur l'analyse des transactions réalisées par les voyageurs durant l'année 2013 (YUWA et DESMOND, 2013).

5.2.2 Annotation du corpus

Compte tenu du volume du corpus collecté, il a été nécessaire de réaliser et mettre en place un système d'annotation d'images et de pages web afin d'extraire un sous-corpus exploitable dans de bonnes conditions.

Pour construire cette vérité terrain, nous avons été assistés par trois annotateurs, qui devaient d'abord déterminer si oui ou non une photographie était une représentation d'une scène de rue dans laquelle une activité humaine pouvait être observable. Ce système a permis l'annotation d'environ 12,8% du corpus. Ce sous-corpus comprend 18 000 images annotées simultanément par 3 annotateurs durant l'été 2014, ce qui représente un total de 55 737 annotations. Sur ces 18 000 images, 6 947 (soit 38,6%) ont été identifiées comme étant des images de rues par au moins un annotateur. Autrement dit, environ 40% des images retournées lors de l'usage des requêtes temporelles associant villes et années, sont des vues de rues.

Les pages web associées aux images visuellement pertinentes ont ensuite pu être annotées grâce à la sélection de l'information textuelle liée à la datation contenue dans ces pages. À partir de chaque image annotée comme étant une vue de rues, les annotateurs ont été invités à identifier séparément l'année correcte de prise de vue de celle-ci. En effet, l'année proposée dans la requête n'est pas toujours l'année exacte de la photographie. Les 4 chiffres censés représenter l'année peuvent, par exemple, représenter la taille de l'image exprimée en pixels, ou encore être une année qui n'est tout simplement pas l'année pendant laquelle la photographie a été prise. Afin de lever ce doute nous avons mis en place un système d'annotation permettant de sélectionner la date de prise de vue de la photographie ainsi que sa provenance, affichant pour aider les annotateurs, en plus de la page web plusieurs informations tels que le titre de la page, le nom du fichier image dans la page web, *etc.* . Le tableau 5.1 donne un aperçu des informations utilisées. Nous pouvons observer dans la deuxième partie du tableau 5.1 que les annotateurs ont trouvé les informations temporelles surtout dans le contenu de la page web. Cela donne un indice quant au potentiel de cette information. Il confirme l'idée qu'une utilisation appropriée de ces données et des connaissances associées peut conduire à de bons résultats en datation automatique de photographies. Nous avons ainsi obtenu une information de première qualité concernant la datation de photographies dans le contexte des pages web.

Remarquons que les métadonnées de photographies sont généralement insuffisantes. Par exemple, l'information temporelle des métadonnées EXIF sur notre jeu de données était disponible uniquement pour 25% des images, et la plupart du temps ces données sont inexactes car seules des dates récentes ont été retrouvées. Des statistiques détaillées concernant ces informations sont disponibles en annexe D à la page 171.

Jeu de données et annotations	
Villes (toutes régions confondues)	50
Années couvertes [de...à...]	191 [1820, 2010]
Annotations réalisées	55 737
Annotations en tant que vues urbaines pour les 3 annotateurs	15 944
Nombre d'images annotées comme étant urbaines par au moins 1 annotateur	6 947
Source des informations temporelles	
Dans la page web	4 071 (25.53%)
Dans le nom du fichier	1 629 (10.22%)
Dans le contenu alternatif de l'image	76 (0.48%)
Dans le titre de la page	69 (0.43%)
Dans le titre de l'image	46 (0.29%)
Autres cas : page introuvable, erreurs, etc.	10 053 (63.05%)

TABLEAU 5.1: Statistiques du jeu de données.

5.2.3 Élaboration d'un sous-corpus de qualité : web-photo 1900-2000

Compte tenu de la variété de la provenance des dates sélectionnées par les annotateurs, nous avons dû vérifier la validité des données. Cette étape a pour objectif, la normalisation des informations de nos annotateurs, afin de procéder à la sélection des documents et de constituer un corpus homogène et de haute qualité. Nous avons dû rejeter les documents pour lesquels les annotations étaient incomplètes, floues, fausses, inconsistantes ou inexploitable. Ceci a restreint la taille du jeu de données mais, à l'inverse, a augmenté la qualité de la vérité terrain, et donc la qualité du jeu de données. Après ce filtrage, il nous restait un total de 3 537 documents.

À partir de ce moment, nous avons décidé de poursuivre la sélection. Afin de nous placer dans un contexte temporel consistant et d'intérêt, nous avons choisi de sélectionner les documents allant des années 1900 à 1999 inclus. Cela représente 2 528 documents, dont il résulte un ensemble de 100 classes correspondant chacune à une année. Il est bien sûr possible de faire varier cette granularité de manière consistante pour de futures études (en l'augmentant).

En outre, nous avons conservé les documents pour lesquels :

- la position de l'image peut être déterminée dans la page web (certains documents en contiennent plusieurs et nous avons jugé utile de pouvoir l'identifier) ;
- au moins une date peut être trouvée grâce à une expression régulière, notre objectif étant de chercher à déterminer si une telle date est représentative de la période de prise de vue de la photographie ;

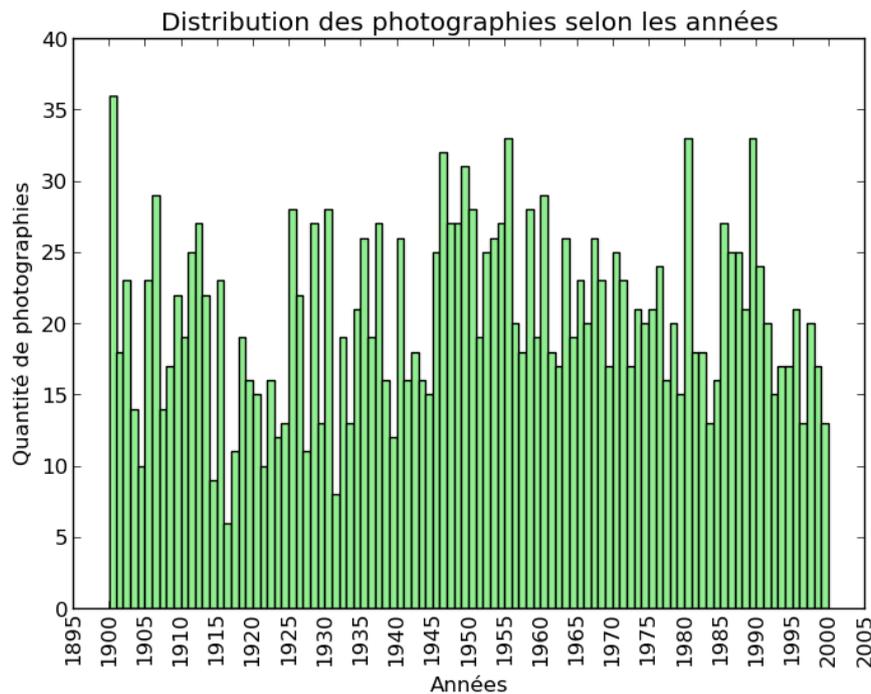


ILLUSTRATION 5.4: Distribution du nombre d'images par année dans notre jeu de données *web-photo* 1900-2000.

histogramme (illustration 5.7) indique la distribution de la résolution des photographies dans notre base.

5.3 Conclusion

Nous avons créé le jeu de données nommé *web-photo* 1900-2000. Ce jeu de données est réparti équitablement dans le temps. Il contient des photographies de rues et des pages web disposant d'un contexte textuel. Compte tenu des caractéristiques que nous avons évoquées dans la section 5.2.3, ce jeu de données constitue un corpus d'évaluation de la datation de photographies, offrant la possibilité d'exploiter le contenu textuel et visuel. Ce corpus comble donc un manque dans la communauté de la recherche d'information. En effet, il s'agit du premier jeu de données permettant d'évaluer la datation de photographies grâce au texte. De plus, ce corpus peut être vu dans la communauté de la vision par ordinateur comme une amélioration du corpus de PALERMO, HAYS et EFROS (2012), puisqu'il couvre un ensemble de 100 classes au grain de l'année, au lieu de 5 classes au grain de la décennie. Il dispose de plus de 2 000 photographies, aussi bien en noir et blanc qu'en couleur.

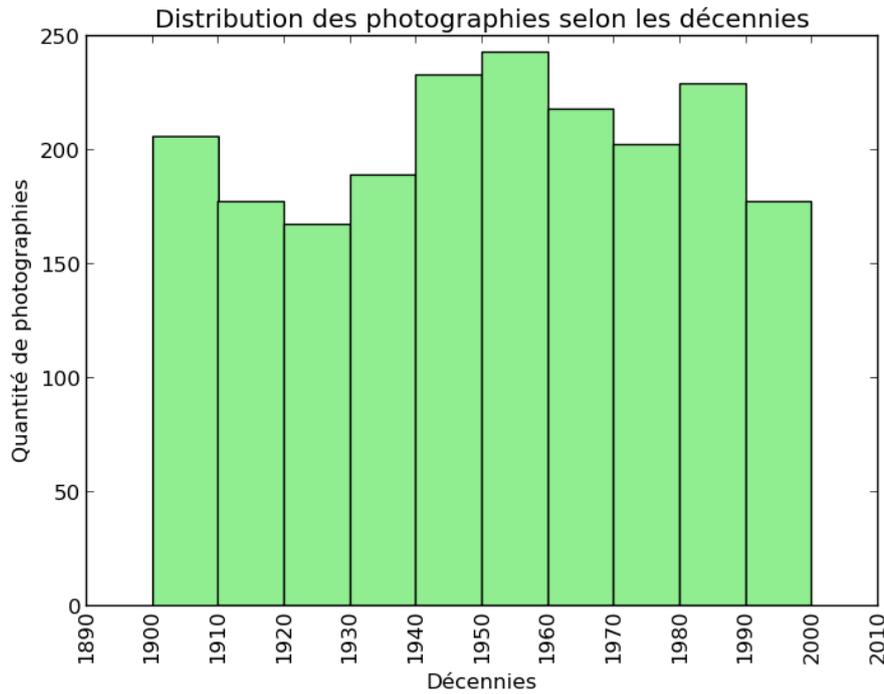


ILLUSTRATION 5.5: Distribution du nombre d'images par décennie dans notre jeu de données *web-photo* 1900-2000.

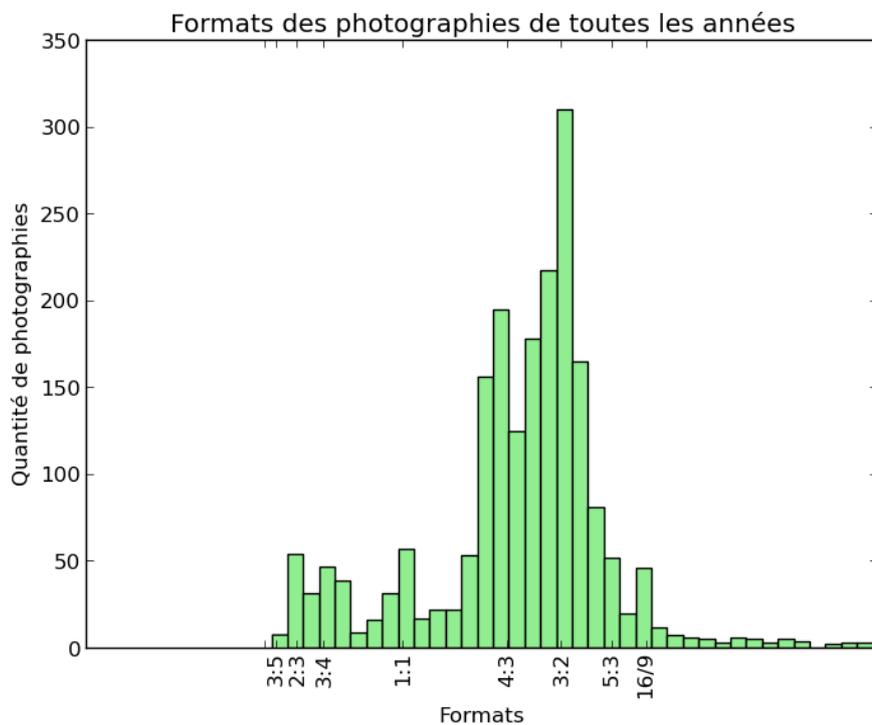


ILLUSTRATION 5.6: Distribution du nombre d'images en fonction de leurs formats dans notre jeu de données *web-photo* 1900-2000.

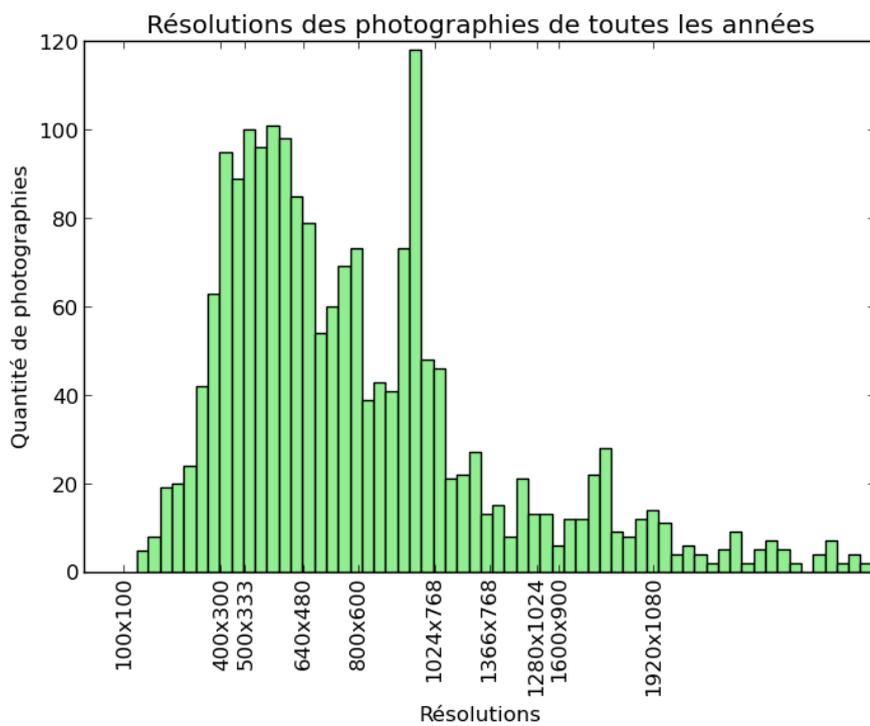


ILLUSTRATION 5.7: Distribution du nombre d'images en fonction de leurs résolutions dans notre jeu de données *web-photo* 1900-2000.

Chapitre 6

Datation visuelle

Sommaire

6.1 Introduction	90
6.2 Nouvelle approche de datation visuelle	91
6.3 Vers une nouvelle méthode de classification ordinale : <i>Avant contre</i> <i>Après</i>	93
6.4 Expérimentations	99
6.5 Conclusion	102

Nous présentons dans ce chapitre une méthode de datation de photographies par l’usage du contenu visuel de celles-ci.

Pour ce travail, nous nous sommes inspirés des travaux de PALERMO, HAYS et EFROS (2012)¹. Nous avons amélioré la méthode de classification utilisée dans leurs travaux en surmontant une limite intrinsèque de leur approche. En effet, ils considèrent la datation d’images comme un problème de classification multi-classes, pour lequel une classe représente un ensemble d’années, mais ignorant l’ordre relatif sous-jacent à l’information temporelle. Dans leur approche soit une prédiction est bonne² soit elle est mauvaise³ mais aucune différence n’est faite entre se tromper d’une décennie ou se tromper de plusieurs.

Nos travaux, s’appuient sur des avancées récentes en classification ordinale (FRANK et HALL, 2001 ; CARDOSO et COSTA, 2007 ; LI et LIN, 2007 ; CHANG, CHEN et HUNG, 2011). Nous considérons les dates comme des attributs à la fois ordonnés et relatifs et nous proposons un cadre spécifique pour les manipuler.

1. En parallèle de nos travaux (MARTIN, 2013 ; MARTIN, DOUCET et JURIE, 2014) d’autres travaux ont été proposés concernant les descripteurs d’images appliqués à la datation d’images couleurs (BERENQUER ARROYO et VIDAL, 2013 ; FERNANDO, MUSELET, KHAN et TUYTELAARS, 2014). Nous apportons nos remarques en conclusion, à la fin de ce chapitre.

2. Période valide.
3. Période invalide.

6.1 Introduction

La méthode que nous avons développée se place dans le cadre des méthodes d'apprentissage statistiques supervisées. Cela signifie qu'il sera nécessaire d'apprendre un modèle, à partir d'exemples d'entraînement annotés et représentatifs du problème. L'objectif du modèle est de prédire des classes. Il est ensuite nécessaire d'analyser l'efficacité du modèle sur de nouvelles données de test pour lesquelles les annotations sont connues et de ce fait indiqueront de la qualité du modèle.

Nous allons commencer par introduire la chaîne de traitement traditionnellement employée pour répondre aux problèmes d'apprentissage statistiques supervisés. La succession des étapes suit généralement un ordre spécifique tel que celui présenté dans l'illustration 6.1.

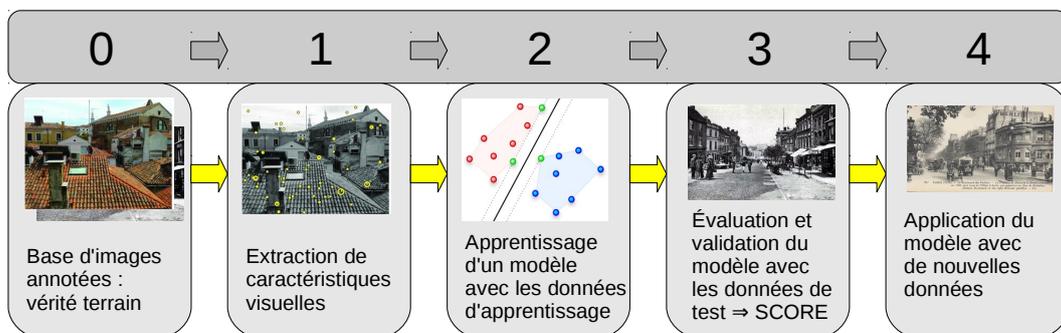


ILLUSTRATION 6.1: Chaîne de traitement en apprentissage statistique supervisé : l'exemple de la classification d'images.

Les étapes de la chaîne de traitement en apprentissage statistique supervisé sont les suivantes :

0 – Données annotées : Cette étape est fondamentale, c'est le pilier de l'apprentissage supervisé. Pour pouvoir se placer dans le cadre des méthodes supervisées, il est absolument nécessaire de disposer de bases de données comportant des annotations similaires à l'objectif auquel nous souhaitons répondre. Dans le cas de la classification d'images, il s'agit d'images annotées représentant l'ensemble des classes à reconnaître.

- 1 – Représentation des données :** Les données à classer, ici les images, doivent être représentées par des attributs numériques qui vont permettre aux algorithmes de les traiter. Cette représentation, appelée aussi signature visuelle ou descripteur visuel dans le cas des images, vise à décrire ces données pour pouvoir les comparer, les dissocier entre elles et traduire au mieux l'objectif visé. Les images appartenant à la même classe doivent avoir des représentations proches voire similaires. Au contraire, les images appartenant à des classes opposées doivent être différenciables selon ces mêmes représentations. L'extraction de caractéristiques suivant une représentation particulière est donc une étape ayant un rôle majeur pour la classification.
- 2 – Apprentissage statistique supervisé :** L'objectif suivant est d'apprendre un modèle à partir de représentations des données. Pour cela, il est nécessaire d'effectuer une association entre les labels et l'espace de représentation pour un ensemble de données annotées ; nous appelons cet ensemble : « l'ensemble d'apprentissage ou d'entraînement ». Cette association permet la calibration des méthodes d'apprentissage, cette calibration se fait sous contraintes de critères à optimiser. Les critères sont choisis en fonction de l'objectif général visé. Pour la classification cela peut être une minimisation des erreurs de prédiction des labels.
- 3 – Évaluation et validation des méthodes :** Une fois le modèle appris sur les données d'apprentissage, il est utilisé sur de nouvelles données de test, non utilisées pour l'entraînement, mais pour lesquelles les classes à prédire sont connues. Cette étape sert à évaluer la qualité du modèle appris. L'évaluation est l'observation de la capacité d'un modèle à retrouver les classes associées aux caractéristiques d'une nouvelle donnée. Le but du modèle appris est donc de prédire au mieux les classes de nouvelles données ; au mieux, signifiant ici, selon un critère d'évaluation traduisant une fois de plus l'objectif visé.
- 4 – Application des méthodes :** Une fois les méthodes évaluées et validées, si les scores apportent une amélioration vis-à-vis des techniques existantes, il est possible d'en faire une utilisation sur de nouvelles données pour lesquelles il n'y a pas de vérité terrain.

6.2 Nouvelle approche de datation visuelle

Le problème que nous abordons dans ce chapitre peut être formulé de la manière suivante : comment prédire la période de réalisation d'une photographie, en utilisant uniquement ses informations visuelles ? Nous supposons que les dates de prises de vues possibles sont réparties en un nombre fini de classes. Ainsi chaque classe représente une période temporelle. Avec une telle définition, nous nous plaçons dans le cadre d'un problème de classification.

Ce problème a récemment été adressé en temps que problème de classification multi-classes standard e.g. au moyen de la stratégie de classification multi-classes *Un contre Un* (PALERMO, HAYS et EFROS, 2012 ; DIAS, MORENO, JATOWT et CAMPOS, 2012). L'une des principales limitations d'une telle stratégie est que la nature ordinale des données n'est pas prise en compte. En effet, dans cette stratégie, toutes les classes sont opposées entre elles sans prendre en compte l'ordre inhérent au problème de la datation. Nous avons présenté les stratégies de décomposition les plus usitées en sous-problèmes binaires afin de traiter des problèmes multi-classes dans la sous-section 4.2.3 à la page 51.

Nous rappelons, que l'entraînement d'un classifieur est basé sur la minimisation d'une fonction de coût. Dans leur approche, le coût produit par une mauvaise classification est le même, quelque soit l'erreur commise. Il n'y a que deux issues, soit la prédiction est valide soit elle ne l'est pas. Pour que la prédiction soit valide il n'y a qu'une seule configuration possible en revanche toutes les autres configurations entraîneront une prédiction invalide. Ainsi une erreur de prédiction de 40 années n'est pas plus pénalisée qu'une erreur de 10 années. Enfin, une autre limitation vient du fait que chaque classifieur est entraîné en utilisant seulement une petite portion des données d'entraînement. Ceci provient de la nature de la stratégie *Un contre Un* (KNERR, PERSONNAZ et DREYFUS, 1990) dans laquelle les données utilisées par chacun des sous-classifieurs ne concerne que les 2 classes à opposer⁴.

Notre méthode s'inspire de l'approche de classification ordinale de FRANK et HALL (2001) que nous avons introduite puis explicitée dans la sous-section 4.2.3 et, de manière plus précise, elle s'apparente aux travaux réalisés par CHANG, CHEN et HUNG (2011).

De ce fait, nous supposons, à l'inverse, qu'il est plus facile de répondre à la question : « Cette photographie a-t-elle été prise avant ou après cette date ? » que de prédire directement la date, ce qui reviendrait à répondre directement à la question « Quelle est la date de prise de vue de cette photographie ? ». Notre raisonnement porte sur le fait que les technologies photographiques ainsi que les contenus des images ont évolué à travers le temps⁵ et peuvent être caractérisés par des jalons. Par exemple, l'arrivée de la pellicule couleur *Kodachrome* en 1935⁶. Par conséquent, nous proposons de formuler la classification temporelle comme une combinaison de classifieurs binaires appris selon une succession de séparations temporelles en utilisant la stratégie *Avant contre Après* que nous avons explicitée dans la sous-section 4.2.3. Cette stratégie de combinaison a déjà été utilisée par le passé dans le cadre de la classification ordinale par FRANK et HALL (2001). Les auteurs ont montré comment transformer tout problème de classification ordinale à K classes en un ensemble simplifié de $K - 1$ sous-problèmes binaires. L'avantage de cette stratégie est que ces sous-problèmes traduisent ainsi l'information ordinale entre les classes.

4. Voir les explications dans la sous-section 4.2.3.

5. Voir le chapitre 2.

6. Voir l'annexe A.2 à la page 158.

En plus de prendre en compte l'information temporelle de manière ordinale, la stratégie *Avant contre Après* permet d'utiliser l'intégralité du jeu de données d'apprentissage pour l'entraînement des $K - 1$ classifieurs binaires. Le nombre de sous-classifieurs binaires à apprendre est, en outre, moindre car seulement $K - 1$ classifieurs doivent être formés, là où la stratégie *Un contre Un* exige la formation de $\binom{K}{2}$ classifieurs (KNERR, PERSONNAZ et DREYFUS, 1990).

Comme nous l'avons énoncé dans la sous-section 4.2.3, nous rappelons également que l'utilisation d'une stratégie *Un contre le Reste* (VAPNIK, 1995) va produire elle aussi un petit nombre de classifieurs (K), mais tous les sous-classifieurs appris vont souffrir d'un fort déséquilibre entre les classes positives et négatives.

6.3 Vers une nouvelle méthode de classification ordinale : *Avant contre Après*

L'usage de la stratégie *Avant contre Après* comporte trois phases. La première est le ré-étiquetage des classes pour l'apprentissage. La seconde concerne la phase d'apprentissage des sous-classifieurs binaires ré-étiquetés. La troisième phase étant la prédiction de la classe temporelle la plus probable pour une image donnée. Cette prédiction se fait à partir de l'ensemble des sous-classifieurs binaires préalablement appris.

Nous considérons les classes temporelles comme étant un ensemble discret fini d'étiquettes. Les étiquettes des classes temporelles sont désignées par $L = \{1 < \dots < k < \dots < K\}$ où K est le nombre de classes temporelles. Ainsi, lors de l'apprentissage, une image sera donnée comme une paire (x_i, y_i) , où $x_i \in \mathbb{R}^d$ représente un vecteur de caractéristiques visuelles à d dimensions de l'image, appelé aussi signature ou descripteur de l'image, et $y_i \in L$ sa classe temporelle.

En se basant sur cette définition, nous allons présenter ces étapes dans les sous-sections suivantes en illustrant les particularités majeures de cette approche au fil de la discussion. Afin d'illustrer notre méthode de manière cohérente avec la section 6.4 nous nous plaçons le même problème de datation de photographies dans lequel l'objectif est d'attribuer la bonne décennie à un cliché parmi 5 décennies allant des années 1930 aux années 1970. Dans cet exemple, les étiquettes $k \in L$ avec $L = \{1 < 2 < 3 < 4 < 5\}$ représentent respectivement les 5 décennies suivantes : *les années 1930*, *les années 1940*, *les années 1950*, *les années 1960* et *les années 1970*.

6.3.1 Ré-étiquetage des sous-classifieurs binaires

Nous définissons un ensemble de $K - 1$ classifieurs binaires, où chaque classifieur Ψ_n a pour rôle de prédire, si, oui ou non, une image appartient à une classe temporelle supérieure à n . Ici $n \in [1, 2, 3, 4]$ représente les années par lesquelles la séparation temporelle des classes d'apprentissage a lieu. C'est à dire dans notre exemple, les années 1940⁷, 1950, 1960 et 1970. L'illustration de cette première étape de ré-étiquetage⁸ est présentée dans le tableau 6.1.

Classifieurs $\Psi_n \setminus$ Étiquettes $k \in L$	1	2	3	4	5
Ψ_4 : Avant contre Après 1970	-	-	-	-	+
Ψ_3 : Avant contre Après 1960	-	-	-	+	+
Ψ_2 : Avant contre Après 1950	-	-	+	+	+
Ψ_1 : Avant contre Après 1940	-	+	+	+	+

TABLEAU 6.1: Ré-étiquetage des exemples d'apprentissage des sous-classifieurs avec la stratégie Avant contre Après.

Ainsi, à partir des 5 classes de notre exemple, 4 sous-classifieurs issus de la stratégie de ré-étiquetage *Avant contre Après* peuvent être appris aux différents seuils de séparation donnés par n .

6.3.2 Apprentissage des sous-classifieurs binaires

Avec la stratégie *Avant contre Après* chacun des sous-classifieurs est calibré en utilisant l'ensemble des données d'apprentissage. La quantité d'exemples utilisables est donc constante pour tous les sous-classifieurs que le jeu de données soit équilibré ou non. Cependant la proportion d'instances positives et négatives varie selon chaque sous-classifieur comme nous l'illustrons sur notre exemple dans le tableau 6.2.

Classifieurs $\Psi_n \setminus$ Étiquettes $k \in L$	1	2	3	4	5
Ψ_4 : Avant contre Après 1970	860				215
Ψ_3 : Avant contre Après 1960	645			430	
Ψ_2 : Avant contre Après 1950	430		645		
Ψ_1 : Avant contre Après 1940	215	860			

TABLEAU 6.2: Impact du ré-étiquetage sur la balance des exemples d'apprentissage des sous-classifieurs Avant contre Après.

7. $n = 1$ représente la séparation de part et d'autre des années 1940, ainsi la classe négative est représentée par la décennie 1930 et la classe positive est représentée par les décennies 1940, 1950, 1960 et 1970.

8. Nous avons détaillé ce processus de ré-étiquetage dans la sous-section 4.2.3.

Dans cet exemple, le jeu de données d'apprentissage est équilibré et comporte 215 exemples d'apprentissage par classe. Ce qui fait un total de 1 075 exemples qui sont utilisés par chacun des sous-classifieurs. Nous pouvons observer dans le tableau 6.2 que le sous-classifieur binaire Ψ_1 appris au début de l'ensemble des classes temporelles utilisera plus d'exemples « après » le seuil de séparation temporelle qu'« avant ». Pour faire face à ce déséquilibre, nous pondérons les classes d'entraînement ré-étiquetées de ces sous-classifieurs par rapport au nombre de classes temporelles originalement présentes dans les parties « avant » et « après ». Nous utilisons des techniques de pondération issues de l'apprentissage sensible aux coûts⁹, telles que celles développées par ELKAN (2001). Nous présentons le calcul des poids des classes ré-étiquetées dans le tableau 6.3.

Classifieurs $\Psi_n \setminus$ Étiquettes $k \in L$	1	2	3	4	5
$\Psi_4 : \text{Avant contre Après } 1970$		1/4			1
$\Psi_3 : \text{Avant contre Après } 1960$		1/3		1/2	
$\Psi_2 : \text{Avant contre Après } 1950$		1/2		1/3	
$\Psi_1 : \text{Avant contre Après } 1940$	1		1/4		

TABLEAU 6.3: Calculs des poids appliqués aux classes ré-étiquetées des sous-classifieurs obtenus avec la stratégie Avant contre Après.

De plus, il est à noter que les exemples distants du seuil de séparation temporelle sont des exemples plus représentatifs de la période de temps concernée. À l'inverse, ceux proches de la séparation sont moins informatifs.

Pour prendre en compte cette spécificité, nous proposons de donner aux exemples d'apprentissage du modèle une importance qui soit relative à leur distance par rapport à la séparation concernée pour chaque sous-classifieur. La solution, que nous proposons, est de donner des poids plus importants aux exemples lointains de la séparation temporelle, car les erreurs commises sur des exemples plus proches des seuils de date sont plus tolérables.

A titre d'exemple, si nous considérons le classifieur Ψ_n , les poids, que nous noterons, $w(n, k)$ sont appliqués aux exemples d'apprentissage en fonction de la classe originale k associée aux exemples de cette manière :

$$w(n, k) = \begin{cases} |n - k| & \text{si } n < k, \\ |n - (k - 1)| & \text{sinon.} \end{cases} \quad (6.1)$$

Nous illustrons le calcul des poids des instances originales dans le tableau 6.4.

9. En anglais : *Cost-sensitive learning*

Classifieurs $\Psi_n \setminus$ Étiquettes $k \in L$	1	2	3	4	5
Ψ_4 : Avant contre Après 1970	4	3	2	1	1
Ψ_3 : Avant contre Après 1960	3	2	1	1	2
Ψ_2 : Avant contre Après 1950	2	1	1	2	3
Ψ_1 : Avant contre Après 1940	1	1	2	3	4

TABLEAU 6.4: Pondérations des exemples d'apprentissage par le poids $w(n, k)$ de l'équation 6.1.

D'autres fonctions de pondération sont possibles, par exemple, le score cumulatif (CHANG, CHEN et HUNG, 2011) ou encore le carré des distances, *etc.* En pratique, elles donnent toutes des performances voisines, pour les expériences que nous avons menées, ce qui peut être dû au faible nombre de classes considérées.

Dans nos expériences pratiques, chaque sous-classifieur appris, avec la stratégie *Avant contre Après*, est un classifieur SVM¹⁰ linéaire (VAPNIK, 1999). Dans l'étape suivante, la prédiction de la classe temporelle finale est calculée par un mécanisme probabiliste combinant chacune des probabilités obtenues par tous les sous-classifieurs appris avec la stratégie *Avant contre Après*.

Nous devons donc introduire une étape supplémentaire pour transformer les sorties des SVMs en probabilités. Pour l'instance x_i , la probabilité donnée par le classifieur Ψ appris à la séparation n sera noté $P(\Psi_n|x_i)$. Nous utilisons le modèle de PLATT (2000) pour calculer¹¹ $P(\Psi_n|x_i)$, la probabilité que l'image x_i appartienne aux classes situées après la séparation temporelle de la date n .

6.3.3 Prédiction à partir des sous-classifieurs binaires

Dans cette troisième étape, les prédictions données par les $K - 1$ sous-classifieurs sont combinées pour obtenir les prédictions de classes temporelles.

Nous avons détaillé dans la sous-section 4.2.3 à la page 51 la stratégie de ré-étiquetage des classes mais aussi les techniques permettant d'obtenir les prédictions des classes ordinales. Nous avons notamment présenté les formules 4.7 (FRANK et HALL, 2001) page 55 et 4.8 (CARDOSO et COSTA, 2007) page 55 qui sont deux techniques de combinaisons des probabilités obtenues par les sous-classifieurs appris avec la stratégie *Avant contre Après*.

Les remarques de CARDOSO et COSTA (2007) concernant les probabilités négatives qui peuvent être produites avec l'approche de FRANK et HALL (2001) sont pertinentes et

10. Séparateur à Vaste Marge

11. Utilisation de la fonction *sigmoid* pour transformer les distances à l'hyperplan séparateur en probabilités.

justifiées. Malheureusement dans leur correction¹² du modèle (voir la formule 6.2 à la page 97) il reste une amélioration possible. En effet, comme nous pouvons le voir dans l'équation 6.2, les probabilités obtenues avec leur modèle ne sont pas « consistantes » pour toutes les classes.

$$\begin{aligned}
 P(y_1|x_j) &= 1 - P(Cl_1|x_j) \\
 P(y_i|x_j) &= (1 - P(Cl_i|x_j)) \prod_{n=1}^{i-1} (P(Cl_n|x_j)), 1 < i < K \\
 P(y_K|x_j) &= \prod_{n=1}^{K-1} (P(Cl_n|x_j))
 \end{aligned} \tag{6.2}$$

Pour les classes $i \in [1, K]$ il n'y a que i classifieurs qui participent au vote final. Une amélioration serait de considérer tous les classifieurs pour toutes les classes.

Ainsi, pour garantir la cohérence de part et d'autre de la séparation temporelle sur l'ensemble des classes, le calcul des probabilités par classe pourrait être fait de cette manière :

$$\begin{aligned}
 P(y_1|x_j) &= \prod_{m=1}^{K-1} (1 - P(Cl_m|x_j)) \\
 P(y_i|x_j) &= \prod_{n=1}^{i-1} (P(Cl_n|x_j)) \prod_{m=i}^{K-1} (1 - P(Cl_m|x_j)), 1 < i < K \\
 P(y_K|x_j) &= \prod_{n=1}^{K-1} (P(Cl_n|x_j))
 \end{aligned} \tag{6.3}$$

Dans la formule 6.3 nous proposons une version améliorée de la formule 6.2 de CARDOSO et COSTA (2007). Ainsi, grâce à notre formulation, pour chacune des classes tous les classifieurs participeront au vote. Nous pensons que cette information est alors plus cohérente.

Décomposons ce qui se produit avec ce modèle de combinaison pour la classe i :

- pour tous les classifieurs avant i , nous prenons les probabilités qu'ils votent après i ;
- et inversement pour tous les classifieurs après i , nous prenons les probabilités qu'ils votent avant i .

Comme pour les autres stratégies, la classe estimée pour l'instance x_j est la classe (y_i) ayant la probabilité la plus élevée.

12. Par souci de commodité, nous rappelons la formule 4.8 (CARDOSO et COSTA, 2007) que nous avons introduite initialement à la page 55 dans la formule 6.2 à la page 97. Cependant, les notations employées dans le chapitre 4 sont dépendantes de ce chapitre.

Ré-écrivons à présent notre nouvelle formulation (6.3) avec les notations du chapitre actuel dans la formule 6.4.

$$\begin{aligned}
 P(1|x_i) &= \prod_{m=1}^{K-1} (1 - P(\Psi_m|x_i)) \\
 P(k|x_i) &= \prod_{n=1}^{k-1} (P(\Psi_n|x_i)) \prod_{m=k}^{K-1} (1 - P(\Psi_m|x_i)), 1 < k < K \\
 P(K|x_i) &= \prod_{n=1}^{K-1} (P(\Psi_n|x_i))
 \end{aligned} \tag{6.4}$$

Nous pouvons simplifier la formule 6.4 en une seule expression.

$$P(k|x_i) = \prod_{n=1}^{k-1} (P(\Psi_n|x_i)) \prod_{m=k}^{K-1} (1 - P(\Psi_m|x_i)) \tag{6.5}$$

Ce modèle donne la probabilité qu'une image x_i appartienne aux classes $k \in L$. Ceci représente donc un nouveau et simple modèle probabiliste de combinaison de classifieurs binaires appris avec la stratégie *Avant contre Après*.

Enfin, la classe de l'image x_i est calculée en maximisant $P(k|x_i) \forall k \in L$, i.e. :

$$\hat{y}_i = \hat{y}(x_i) = \arg \max_{k \in L} P(k|x_i). \tag{6.6}$$

Nous pouvons appliquer la formule 6.5 sur notre exemple illustratif. Nous pouvons observer le détail de la combinaison des probabilités de chacune des 5 classes à partir des 4 classifieurs issus de la stratégie *Avant contre Après* ci dessous :

$$\begin{aligned}
 P(5|x_1) &= P(\Psi_1|x_1) * P(\Psi_2|x_1) * P(\Psi_3|x_1) * P(\Psi_4|x_1) \\
 P(4|x_1) &= P(\Psi_1|x_1) * P(\Psi_2|x_1) * P(\Psi_3|x_1) * (1 - P(\Psi_4|x_1)) \\
 P(3|x_1) &= P(\Psi_1|x_1) * P(\Psi_2|x_1) * (1 - P(\Psi_3|x_1)) * (1 - P(\Psi_4|x_1)) \\
 P(2|x_1) &= P(\Psi_1|x_1) * (1 - P(\Psi_2|x_1)) * (1 - P(\Psi_3|x_1)) * (1 - P(\Psi_4|x_1)) \\
 P(1|x_1) &= (1 - P(\Psi_1|x_1)) * (1 - P(\Psi_2|x_1)) * (1 - P(\Psi_3|x_1)) * (1 - P(\Psi_4|x_1))
 \end{aligned}$$

Comme dit précédemment, la décennie qui sera attribuée à la photographie sera celle recevant la probabilité maximale (6.6).

Bien qu'étant différente, cette formulation est liée à celle présentée par FRANK et HALL (2001) et CARDOSO et COSTA (2007). L'avantage et la différence majeure de notre approche résident dans le fait que pour chacune des classes tous les classifieurs sont utilisés dans la phase de combinaison. Le modèle ainsi appris, incorpore plus d'informations pour calculer la prédiction à l'inverse des autres stratégies.

L'expression que nous proposons dans l'équation 6.5 est intuitive ; elle tire profit de l'ensemble de tous les classifieurs produisant ainsi des prédictions plus cohérentes et plus complètes. Les méthodes ordinales existantes effectuant des combinaisons de probabilités telles que celles de FRANK et HALL (2001) ou CARDOSO et COSTA (2007) utilisent respectivement la différence entre les probabilités de deux classifieurs ou le produit d'une partie des probabilités des classifieurs (les détails de ces méthodes sont donnés dans la sous-section 4.2.3). À l'inverse, notre méthode propose une combinaison plus complète car elle exploite l'information d'intersection entre tous les classifieurs.

En outre, la combinaison proposée est basée sur le produit des probabilités. Ce fonctionnement est préférable à des méthodes telles que celles proposées par LI et LIN (2007) qui opèrent avec la somme des valeurs de confiance seuillées issues des classifieurs. En effet, ces derniers soulèvent un problème vis-à-vis de la statistique puisqu'ils combinent des probabilités en utilisant des sommes et non des produits.

6.4 Expérimentations

Le *framework* de classification ordinaire que nous avons proposé pour la datation des images couleur est validé expérimentalement sur le jeu de données de PALERMO, HAYS et EFROS (2012), qui est un jeu de données de référence pour la datation d'images couleur. Quelques images sont présentées en exemple dans l'illustration 6.2. Le jeu de données et leurs codes sources sont disponibles en ligne¹³. La base de données est constituée de 1 325¹⁴ images couleur pour lesquelles les dates de prises de vues sont connues. Ces images proviennent du site *flickr* et ont été inspectées manuellement afin de supprimer le contenu non photographique et parasite. Elles ont été divisées en 5 classes d'âges correspondant aux 5 décennies de 1930 à 1970. L'ensemble est équilibré, avec un total de 265 images par classe. Pour que notre comparaison soit équitable, nous prenons exactement la même configuration (y compris les sept mêmes caractéristiques visuelles, nous les avons présentées brièvement dans la sous-section 3.2.2 à la page 38) et nous utilisons les codes fournis par les auteurs.

La performance est évaluée en utilisant l'Erreur Absolue Moyenne (*MAE*)¹⁵, qui est définie comme étant l'écart moyen, ici exprimé en décennies, entre la classe prédite et la véritable classe,

$$MAE = \frac{1}{|T|} \sum_{x_i \in T} |y(x_i) - \hat{y}(x_i)|, \quad (6.7)$$

13. <http://graphics.cs.cmu.edu/projects/historicalColor/> dernier accès août 2015

14. Dans l'article, le nombre d'images de leur jeu de données est annoncé à 1 375 par erreur, nous reportons ici le nombre d'images présentes dans l'archive disponible à l'adresse <http://graphics.cs.cmu.edu/projects/historicalColor/> dernier accès août 2015

15. En anglais : *Mean Absolut Error*, également appelée *Ranking loss*



(a) 1930



(b) 1940



(c) 1950



(d) 1960



(e) 1970

ILLUSTRATION 6.2: Exemples d'images de la collection de PALERMO, HAYS et EFROS (2012).

Dans l'équation 6.7 T désigne l'ensemble d'apprentissage, $y(x_i)$ la véritable classe et $\hat{y}(x_i)$ la classe prédite pour une image donnée x_i .

Le MAE (6.7) est l'une des mesures de référence pour évaluer la classification ordinaire (LI et LIN, 2007 ; CARDOSO et COSTA, 2007 ; BACCIANELLA, ESULI et SEBASTIANI, 2009a ; CHANG, CHEN et HUNG, 2011). Bien que la précision¹⁶ moyenne (ACC) ne soit pas très adéquate pour l'évaluation de la datation d'images, car elle ne tient pas compte de l'importance de l'erreur, nous rapportons les résultats obtenus avec celle-ci afin de permettre la comparaison avec les travaux de PALERMO, HAYS et EFROS (2012).

16. En anglais : *Accuracy*

Nous utilisons le même protocole d'évaluation que PALERMO, HAYS et EFROS (2012). Dans chacune des décennies, 215 images sont sélectionnées au hasard et sont utilisées pour l'apprentissage¹⁷ tandis que les 50 images restantes sont utilisées pour les tests. On peut remarquer que, dans la stratégie standard *Un contre Un* sur ce jeu de données, chaque classifieur opposant donc 2 classes à la fois, utilise seulement 430 exemples. Dans notre approche, chaque classifieur exploite cette fois, les 1 075 exemples¹⁸ du jeu de données d'apprentissage. Selon nous, cela permet d'améliorer la qualité de la classification finale.

L'évaluation comparative est effectuée par rapport aux travaux de référence de PALERMO, HAYS et EFROS (2012) pour la datation des images couleur. Nous avons aussi calculé les résultats avec 3 méthodes de classification ordinale de l'état de l'art, à savoir

- le *framework* de réduction de LI et LIN (2007) ;
- le modèle de combinaison probabiliste de FRANK et HALL (2001) ;
- le modèle de combinaison probabiliste de CARDOSO et COSTA (2007).

Les 3 méthodes de combinaison probabilistes (la notre, FRANK et HALL (2001) et CARDOSO et COSTA (2007)) sont basées sur la même stratégie *Avant contre Après* .

La spécificité de notre approche est l'utilisation de techniques de sensibilité aux coûts¹⁹ basés sur le calcul d'un coût absolu, afin de tenir compte de la relation de distance existante entre les classes. Nous rappelons ici que ces techniques ne sont pas utilisées dans les travaux originaux de FRANK et HALL (2001) ni dans ceux de CARDOSO et COSTA (2007).

La *MAE* (et l'*ACC*) de ces 5 approches sur le jeu de données de test sont calculées et moyennées sur 10 découpages aléatoires entre les ensembles d'apprentissage et de test, comme spécifié dans le protocole expérimental de PALERMO, HAYS et EFROS (2012). Nous rapportons également les valeurs d'écart-type. Les résultats correspondants sont présentés dans le tableau 6.5, où il est montré que notre approche fonctionne significativement mieux que celles des concurrents.

Les comparaisons sont effectuées en utilisant les mêmes caractéristiques visuelles et le même protocole expérimental tel que défini dans PALERMO, HAYS et EFROS (2012). Nous rappelons également les résultats qu'ils ont obtenus sur la précision d'annotateurs humains non formés sur cette problématique et ce jeu de données. Nous avons de plus ajouté les résultats obtenus par un tirage aléatoire des classes parmi les 5 disponibles ainsi que la prédiction constante de la classe centrale. À ce propos, nous pouvons observer que bien qu'ayant des *ACC* similaires, le choix de la classe centrale lui, minimise la distance d'erreur moyenne (*MAE*) dans ce problème équilibré en comparaison du choix aléatoire.

17. Dans le document original, la taille de l'ensemble d'apprentissage par classe est indiqué par erreur à 225 images par classe.

18. Comme nous l'avons présenté dans le tableau 6.2 à la page 94.

19. En anglais : *Cost-sensitive*

	ACC (\pm écart-type)	MAE (\pm écart-type)
Classe aléatoire	20	1.6
Valeur centrale : <i>années 1950</i>	20	1.2
Annotateurs humains non formés (PALERMO, HAYS et EFROS, 2012)	26.0 (\pm N/A)	N/A
FRANK et HALL (2001)	41.36 (\pm 1.89) (*)	0.99 (\pm 0.05) (**)
LI et LIN (2007)	35.92(\pm 4.69) (*)	0.96(\pm 0.06) (**)
CARDOSO et COSTA (2007)	41.32 (\pm 2.76)	0.95 (\pm 0.04)(**)
PALERMO, HAYS et EFROS (2012)	44.92(\pm 3.69)	0.93(\pm 0.08) (*)
Approche proposée	42.76 (\pm1.33)	0.87 (\pm0.05)

TABLEAU 6.5: Performance de l'approche proposée sur le jeu de données de PALERMO, HAYS et EFROS (2012), ainsi que les comparaisons avec les approches connexes.

Pour chacun des résultats, nous avons calculé si nos résultats sont statistiquement significatifs en utilisant le « test t »²⁰ sur deux échantillons indépendants. La signification statistique par rapport à notre approche a été systématiquement calculée : (*) veut dire que $p < 0.05$, (**) veut dire que $p < 0.01$. En termes de MAE , cela démontre que notre approche est nettement meilleure que toutes les autres, avec $p < 0,05$ par rapport à PALERMO, HAYS et EFROS (2012), et $p < 0,01$ par rapport aux autres. Fait intéressant, en termes d' ACC , la différence n'est pas statistiquement significative entre notre méthode et celle de PALERMO, HAYS et EFROS (2012) et CARDOSO et COSTA (2007), alors que nous surpassons nettement les deux autres approches. Par conséquent, en termes statistiques, il est juste de dire que notre approche fonctionne mieux concernant la MAE et est à égalité pour la première place en termes d' ACC .

6.5 Conclusion

Nous avons présenté une nouvelle technique qui améliore la datation d'images. Ce problème de classification est particulier en ce sens que ses classes suivent une relation d'ordre. Nous avons redéfini la problématique de la datation de photographies à travers cette question : « Cette photographie a-t-elle été prise avant ou après cette date ? ». Nous avons mis en place un *framework* de classification ordinaire utilisant une stratégie *Avant contre Après*. Nous avons proposé une nouvelle stratégie de combinaisons des résultats des sous-classifieurs binaires obtenus à chaque seuil de séparation temporelle. Cette stratégie utilise les réponses de chaque sous-classifieur.

20. Le test de Student

Grâce à cette combinaison de classifieurs binaires, nous avons tiré parti de la nature ordinaire du temps, comme nous l'avons démontré dans nos expériences. Notre approche a dépassé les techniques de l'état de l'art. Fondamentalement, notre technique minimise la distance d'erreur moyenne, tout en maintenant une précision à la hauteur de l'état de l'art. En d'autres termes, notre approche donne le même nombre de mauvaises classifications, mais en moyenne, celles-ci tombent plus près de la valeur réelle.

À l'avenir, nous prévoyons d'apprendre des années plutôt que des périodes plus larges, dans l'hypothèse que les prédictions plus précises pourront être obtenues. Une autre hypothèse de travail serait de « deviner » à partir des données, des seuils de séparation temporels qui soient caractéristiques de la photographie. De plus, nous pensons également que l'étude de la datation d'image doit bénéficier d'ensembles de données plus larges. C'est à dire, avec plus de classes et plus d'images à la fois noir et blanc et couleur ; de telle manière que l'on puisse mieux exploiter les approches basées sur l'ordre et la distance. C'est pourquoi, nous avons développé un jeu de données de 100 années comportant 2 041 documents numériques composés de photographies de rues et de pages web contextualisantes. Ce jeu de données est présenté dans le chapitre 5 et évalué sur la tâche de la datation textuelle dans le chapitre 7.

Remarques

Nous avons pris connaissance des travaux de FERNANDO, MUSELET, KHAN et TUYTELAARS (2014) parus environ 6 mois après notre publication (MARTIN, DOUCET et JURIE, 2014). Dans leurs travaux, les auteurs proposent de nouveaux descripteurs couleurs, basés sur les méthodes de calibration radiométrique de LIN, GU, YAMAZAKI et SHUM (2004). Ils indiquent que les nouveaux descripteurs couleurs qu'ils proposent, tentent d'exploiter des caractéristiques discriminantes du type d'appareil ayant réalisé le cliché. Les performances de leurs descripteurs ont été prouvées sur les données du jeu de PALERMO, HAYS et EFROS (2012).

Les résultats en terme de MAE ou d'autres mesures de classification ordinaire ne sont pas reportés. Cependant une matrice de confusion est présentée dans leur article pour les meilleurs descripteurs c'est à dire lors de la classification avec un SVM non linéaire en utilisant les descripteurs : $0 - CD$, $1 - CD$, $2 - CD$ ²¹ et CA ²². Nous avons calculé à partir de la matrice de confusion, les performances de leur approche selon des mesures de classification ordinaire²³.

21. « Des histogrammes en 3 dimensions des composantes RVB ($0 - CD$), dérivée première ($1 - CD$) et dérivée seconde ($2 - CD$) » – (FERNANDO, MUSELET, KHAN et TUYTELAARS, 2014).

22. « CA représente les angles entre les couleurs dans l'espace RVB » – (FERNANDO, MUSELET, KHAN et TUYTELAARS, 2014).

23. Ces mesures sont introduites dans la section 4.3 à la page 56.

<i>ACC</i>	<i>MAE</i>	<i>AMAE</i>	<i>TauB</i>
85.48	0.26030	0.26056	0.82

TABLEAU 6.6: Résultats de FERNANDO, MUSELET, KHAN et TUYTELAARS (2014) que nous avons calculés à partir de la matrice de confusion reportée dans leur article.

Les résultats qu'ils obtiennent sont extrêmement bons (voir le tableau 6.6). Cependant, nous avons observé plusieurs erreurs que les auteurs ont commises au regard du protocole expérimental original de PALERMO, HAYS et EFROS (2012). Dans le protocole original, les 10 plis aléatoires produisent un total de 500 images de test par classes. En effet, dans l'article de PALERMO, HAYS et EFROS (2012) les auteurs ont utilisé par classe 50 images pour le test et 215 images pour l'apprentissage. Le nombre d'images pour l'apprentissage à été reporté à tort dans l'article de PALERMO, HAYS et EFROS (2012) comme étant égal à 225. Dans leurs expériences FERNANDO, MUSELET, KHAN et TUYTELAARS (2014) ont donc utilisé par erreur 225 images par classes pour l'apprentissage. Compte tenu du fait qu'il y ait 265 images disponibles par classe ceci n'a laissé que 40 images au lieu de 50 par classe pour le test. Au regard de ces différences techniques dans le processus expérimental, les résultats qu'ils ont obtenus ne sont donc techniquement pas comparables. De plus, dans leurs expériences une erreur dans le nombre d'images testées s'est produite. En effet, on observe pour la deuxième colonne dans la matrice de confusion²⁴ un total de $11 + 326 + 15 + 18 + 20 = 390$ images testées pour cette classe. Le nombre d'images testées dans les autres classes est lui correct (400). Avec cette différence de 10 images cela indique qu'ils ont omis une image de test pour la deuxième classe et ce pour chacun des 10 plis. Cette omission se ressent dans les évaluations entre la *MAE* et la *AMAE* car si le jeu de test avait été correctement équilibré ces deux métriques auraient donné le même score (voir le tableau 6.6).

Les implémentations utilisées pour obtenir leurs nouveaux descripteurs couleurs ne sont pas publiés. Il serait intéressant de pouvoir comparer les performances de leurs descripteurs (*i.e.* dans le cas linéaire une *ACC* de 54.4 ± 4.0 et dans le cas non linéaire une *ACC* de 85.5 ± 1.5) dans le cadre de la classification ordinaire. En effet, le modèle qu'ils utilisent n'est pas un modèle de classification ordinaire.

24. Cette matrice de confusion est présentée dans leur article (FERNANDO, MUSELET, KHAN et TUYTELAARS, 2014). Nous l'avons reportée dans l'annexe I à la page 207. Nous y indiquons les problèmes liés aux typographies présentes dans l'article de PALERMO, HAYS et EFROS (2012) et l'impact sur le protocole expérimental employé.

BERENGUER ARROYO et VIDAL (2013) ont montré que l'usage d'un descripteur couleur bas-niveau basé sur une discrétisation des couleurs transportent des informations utiles pour l'analyse de la temporalité des images. En effet, ce descripteur, *Color Naming* (BENAVENTE, VANRELL et BALDRICH, 2008), utilisé seul, bat déjà les scores de PALERMO, HAYS et EFROS (2012) avec une *ACC* de 50.6%. Une fois combiné avec les descripteurs de PALERMO, HAYS et EFROS (2012) les auteurs ont atteint un score total de 65.84% d'*ACC*. Cependant, nous n'avons pas pu calculer les performances en terme de mesures de classification ordinale. Il serait donc une fois de plus intéressant de tester l'efficacité de ce descripteur pour la datation des photographies dans un cadre de classification ordinale.

La limite que nous avons mise en avant dans ce chapitre concernant la stratégie de classification multi-classes *Un contre Un* est aussi présente dans les travaux de DIAS, MORENO, JATOWT et CAMPOS, 2012. Cependant, les auteurs ont expérimenté un découpage non régulier du temps (voir sous-section 3.2.2 à la page 38). Cette approche semble pertinente compte tenu de l'évolution de la photographie et du découpage du temps « adaptatif » à chaque application que nous avons exprimée dans la sous-section 1.2 à la page 10. En revanche, compte tenu de cette configuration, l'usage de la mesure *MAE* devient particulier dans le sens où son interprétation est dépendante de ce découpage. En effet, cette dernière ne s'exprime plus en terme de moyenne d'erreur de décennies, ce qui peut rendre compliqué l'analyse des expériences.

Comparaison avec d'autres méthodes de classification ordinale

Nous avons testé notre méthode sur une série de 10 jeux de données ordinaux : « *automobile* », « *bondrate* », « *contact-lenses* », « *eucalyptus* », « *newthyroid* », « *pasture* », « *squash-stored* », « *squash-unstored* », « *tae* », « *winequality-red* ». Ces jeux de données ont souvent été utilisés dans la littérature des méthodes de classification ordinale²⁵. Il était important pour nous de nous comparer aux autres méthodes de classification ordinale sur ces jeux de données. Ainsi nous avons pu situer notre méthode de classification ordinale présentée dans ce chapitre²⁶ par rapport à 9 méthodes standard de l'état de l'art qui sont rassemblées et évaluées dans l'article de référence de SÁNCHEZ-MONEDERO, GUTIÉRREZ, ANTONIO, TINO et HERVÁS-MARTÍNEZ (2013).

25. Ces jeux de données proviennent de l'UCI (BACHE et LICHMAN, 2013) <http://archive.ics.uci.edu/ml/> dernier accès août 2015

26. Cette méthode a fait l'objet d'une publication (MARTIN, DOUCET et JURIE, 2014).

Nos résultats ainsi que ceux obtenus dans l'article de SÁNCHEZ-MONEDERO, GUTIÉRREZ, ANTONIO, TINO et HERVÁS-MARTÍNEZ (2013) sont consultables, dans une série de tableaux comparatifs, en annexe J à la page 209. Dans ces tableaux figurent 10 méthodes de classification ordinale comparées à travers 10 jeux de données selon 4 mesures d'évaluation couramment usitées en classification ordinale²⁷ : l'*ACC*, la *MAE*, la *AMAE* et le τ_b , reportés dans les tableaux J.1, J.2, J.3 et J.4.

Les résultats obtenus ne placent pas toujours notre méthode dans les meilleures. Cependant, si l'on observe attentivement ces résultats, nous pouvons remarquer que selon les différents jeux de données l'ensemble des méthodes comparées ont des comportements variables. En effet, chacune des 10 méthodes a ses avantages et ses inconvénients. Il y a certes des méthodes préférentielles comme par exemple la méthode SVR-PCDOC qui est celle développée par les auteurs du comparatif (SÁNCHEZ-MONEDERO, GUTIÉRREZ, ANTONIO, TINO et HERVÁS-MARTÍNEZ, 2013). Notre méthode a l'avantage d'être facile d'accès dans sa compréhension en offrant simultanément une bonne rapidité d'exécution du fait du faible nombre de sous-classifieurs à apprendre. Cela en fait potentiellement une alternative intéressante qui offre des performances proches et comparables des méthodes existantes. Les améliorations futures pourront rendre notre méthode plus performante sur ces données ordinales qui présentent de caractéristiques différentes de celles que nous avons employées dans le reste de ce chapitre.

27. Ces mesures sont introduites dans la section 4.3 à la page 56.

Chapitre 7

Datation textuelle

Sommaire

7.1 Introduction	108
7.2 Démarche conceptuelle	110
7.3 Chaîne de traitement de la réconciliation temporelle	114
7.4 Expériences et résultats	120
7.5 Conclusion	123

La classification temporelle de contenus web nécessite de la « connaissance » à leur sujet. Ceci est particulièrement le cas lorsque le contenu contient plusieurs dates et qu'une « interprétation » humaine est nécessaire afin de choisir la bonne information temporelle. Le défi de la datation devient encore plus complexe lorsque les images doivent être datées sur l'unique base du contenu textuel les décrivant. Dans ce chapitre, nous présentons une approche d'horodatage basée sur la sémantique extraite des documents et nous expliquons le fonctionnement de notre chaîne de traitement de réconciliation temporelle. La réconciliation temporelle¹ est un processus qui tire parti des informations provenant d'entités nommées pour inférer la période temporelle la plus probable pour un document. Le jeu de données que nous utilisons pour nos expérimentations est présenté dans le chapitre 5.

1. Le contenu de ce chapitre a fait l'objet d'une publication (MARTIN, SPANJOL et DOUCET, 2015).

7.1 Introduction

Dans l'ère du web 2.0 et des médias sociaux, un nouveau comportement d'utilisateurs a émergé. Les consommateurs se professionnalisent car non seulement ils consomment mais peu à peu produisent du contenu, ce qui leur vaut le nom de *Prosommateur*. Ils génèrent des contenus numériques à des vitesses très importantes sur un grand nombre de plates-formes de médias sociaux. Généralement, ces contenus intègrent également des images qui proviennent de diverses sources, notamment provenant des appareils photographiques numériques modernes, mais résultant aussi de la numérisation par les scanners. En tant que telles, ces images couvrent différentes périodes de temps et portent une information temporelle dans leurs représentations des lieux, des gens, des événements, des faits, *etc.* C'est sur cette information temporelle que porte notre recherche.

Opening Ceremony Summer Olympics, Los Angeles



Photo credit / source: summergames.ap.org

ILLUSTRATION 7.1: Dans cet exemple aucune mention de date n'apparaît, seule la légende peut donner une information de temporalité : « Cérémonie d'ouverture des Jeux Olympiques d'été de Los Angeles ». (source summergames.ap.org)

Un problème inhérent aux contenus des médias sociaux, et en particulier des images, est que leur date de création n'est pas toujours valide. Bien que les nouvelles caméras et les scanners permettent aux utilisateurs de dater automatiquement le contenu, cette information est en réalité souvent peu fiable puisque les appareils sont la plupart du temps mal configurés. Notamment lors de la numérisation, la date de prise de vue du cliché sera souvent la date de numérisation plutôt que la date réelle de prise de vue car le changement de cette date dépend de la volonté de l'utilisateur et cette manipulation n'est pas forcément accessible à tous (c.f. sous-section 2.2.2). Cela conduit à des contenus qui disposent d'informations temporelles incohérentes et qui ne peuvent par conséquent pas être correctement exploitées. Notre objectif est donc d'améliorer l'exploitation des connaissances temporelles présentes aux documents.

L'illustration 7.1 est un premier exemple d'image accompagnée d'une légende. Dans la légende se trouvent des indices temporels implicites qui permettent d'obtenir des informations de temporalité liées à l'image. Dans cet exemple, la légende mentionne les jeux olympiques d'été de *Los Angeles*. Il est possible de savoir qu'ils ont eu lieu en 1984. Cette information extrêmement précise permet à un amateur de sport de positionner la photographie dans le temps de manière aisée.

L'illustration 7.2 est un second exemple dans lequel plusieurs dates sont présentes. La difficulté ici est de sélectionner la bonne date, sachant que la date de prise de vue ne fait pas nécessairement partie des dates mentionnées dans le texte environnant. Le contexte sémantique permet encore une fois de désambiguïser temporellement cette photographie en invalidant les dates parasites.

Comme présenté dans ces deux exemples, l'objectif de notre approche est d'exploiter les informations sémantiques afin d'obtenir des indices temporels. L'une des techniques permettant d'élever le texte à un niveau sémantique s'appuie sur les entités nommées. Notre démarche repose donc sur l'exploitation des entités nommées et l'extraction de relations temporelles accessibles à partir de ces dernières.

LOS ANGELES 1984 Recherche ac

Ulrike Meyfarth revient pour une deuxième médaille d'or et un nouveau record



10/08/1984
Lorsqu'Ulrike Meyfarth avait comblé ses supporters, à domicile, en obtenant la médaille d'or du saut en hauteur aux Jeux olympiques de 1972 à Munich, elle était entrée du même coup dans les annales en devenant la plus jeune championne olympique jamais sacrée en athlétisme.

ILLUSTRATION 7.2: Dans cet exemple plusieurs mentions de dates apparaissent : 1984, 10/08/1984 et 1972. (source olympic.org)

7.2 Démarche conceptuelle

Notre hypothèse est la suivante : la « véritable » information temporelle qui indique quand a été réalisée une photographie peut se trouver à la fois à l'intérieur de l'image et dans le texte l'entourant quand il existe. Cependant, horodater une image en se basant sur le texte qui l'entoure n'est pas une tâche triviale : la date peut ne pas être explicitement mentionnée, elle peut décrire un élément distinct de la photographie, et il est également possible qu'il y ait plusieurs dates parmi lesquelles choisir. Dans de tels cas, un être humain est généralement capable d'identifier et d'« interpréter » le contenu afin de déterminer la période temporelle réellement concernée. Pour ce faire, l'être humain exploite la sémantique du contenu textuel et applique ses connaissances et ses capacités cognitives pour effectuer un alignement temporel afin de déduire la date vraisemblablement la plus juste.



ILLUSTRATION 7.3: Photographie d'illustration d'un article du 5 octobre 2012 sur le lancement de l'iPhone. (source darkroom.baltimoresun.com)

Nous allons exposer le principe de notre méthode à partir de l'illustration 7.3. Il s'agit d'une capture d'écran d'une page web² dans laquelle figure un article sur Steve Jobs présentant l'iPhone à sa sortie en 2007. Cet article date du 5 octobre 2012. Nous allons voir quelles sont les informations temporelles que la légende de cette photographie peut nous révéler.

Nous lisons sur la légende :

« January 9, 2007 : Apple^a CEO Steve Jobs^b holds up an Apple iPhone^c at the MacWorld Conference^d in San Francisco, California^e. (Paul Sakuma/AP Photo) »

Dans cette légende se trouve une série de termes correspondant à des entités nommées. Pour chacune d'entre elles, il est possible d'obtenir des informations temporelles, comme nous pouvons le voir dans les pages Wikipédia correspondantes à ces entités nommées dans l'illustration 7.4.

Ces ressources nous permettent entre autres d'obtenir des dates de début (création, naissance, etc.) et de fin (décès, etc.), ce qui nous permet de déduire des intervalles correspondant aux différentes entités :

². La page est présentée en langue anglaise car nous avons effectué l'extraction des entités nommées dans cette langue.



ILLUSTRATION 7.4: Informations extraites des entités nommées détectées. (source Wikipedia)

- Apple^a : [1 avril 1976, ∞[;
- Steve Jobs^b : [24 février 1955, 5 octobre 2011];
- iPhone^c : [29 juin 2007, ∞[;
- MacWorld Conference^d : pas de données temporelles exploitables;
- San Francisco, California^e : [29 juin 1776, ∞[.

À partir de ces informations, nous pouvons déterminer une période commune pendant laquelle les intervalles temporels de ces différentes entités nommées se recoupent. La borne inférieure de cet intervalle commun sera ici la date de sortie de l'iPhone, le 29 juin 2007, et la borne supérieure sera la date de la mort de Steve Jobs, le 5 octobre 2011. De cette approche résulte un intervalle de confiance de 4 années, 3 mois et 6 jours. L'illustration 7.5 présente notre approche par le biais d'une frise chronologique.

Ainsi la date de l'article (5 octobre 2012) peut être invalidée, tandis que la date du 9 janvier 2007 peut être conservée grâce au principe de réconciliation temporelle.

a. <https://fr.wikipedia.org/wiki/Apple>, dernier accès août 2015

b. https://fr.wikipedia.org/wiki/Steve_Jobs, dernier accès août 2015

c. <https://fr.wikipedia.org/wiki/IPhone>, dernier accès août 2015

d. https://fr.wikipedia.org/wiki/Macworld_Conference_&_Expo, dernier accès août 2015

e. https://fr.wikipedia.org/wiki/San_Francisco dernier accès août 2015

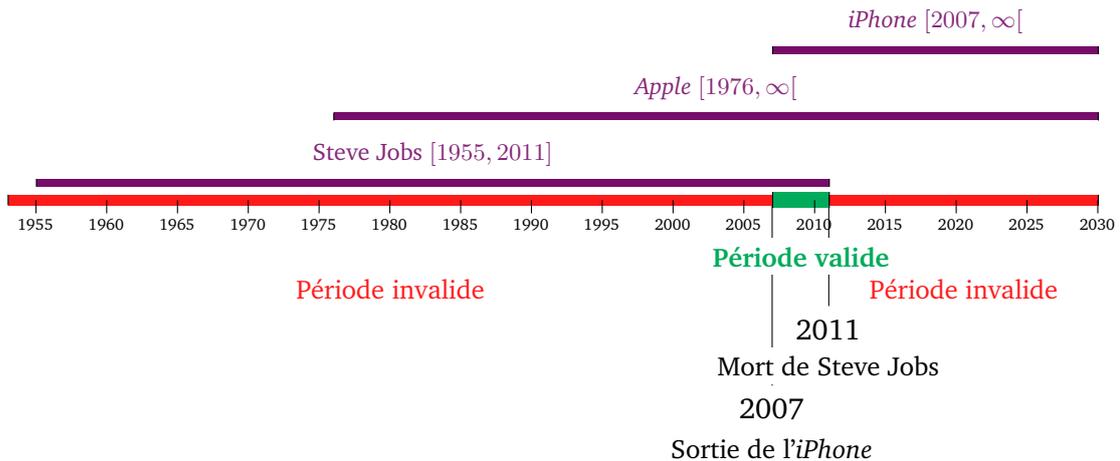


ILLUSTRATION 7.5: Démonstration conceptuelle de l'approche de la réconciliation temporelle.

Remarques et observations

Une limitation de cette approche vient du fait qu'une entité nommée peut venir parasiter le *continuum* temporel et donc invalider le fonctionnement de la réconciliation temporelle. Par ailleurs, l'approche dépend de l'existence de ressources externes décrivant les entités nommées, à la fois pour la détection, la désambiguïsation mais aussi pour l'association avec les faits temporels ; or ces ressources sont parfois indisponibles (e.g. , *Macworld*) ou encore peu pertinentes dans notre cadre de travail (e.g. , *San Francisco*).

D'après nos observations, les moteurs de recherche ne datent pas les photographies mais utilisent des heuristiques associant les images à l'information textuelle la plus proche³, selon les principes habituels de la recherche d'information par mots-clés (BAEZA-YATES, RIBEIRO-NETO et al., 1999). Ainsi, effectuer une recherche de photographies datées par mots-clés en y adjoignant une date ne fonctionnera pas toujours, comme nous venons de le décrire. Cette heuristique fonctionne cependant la plupart du temps, comme chacun peut le vérifier en ligne, et les méthodes des moteurs de recherche commerciaux sont un élément essentiel de l'état de l'art que nous utiliserons dans notre évaluation comparative.

D'autres approches pragmatiques sont envisageables comme par exemple choisir parmi l'ensemble des dates trouvées dans la page web qui respecte un certain critère : la date la plus fréquente, la date moyenne, etc.

3. Cette proximité peut être mesurée par caractère, mots, phrases, blocs, etc. Selon les moteurs.

Il est aussi possible d'agir de manière plus sophistiquée grâce à l'usage de bases de connaissances. En effet, grâce à l'émergence des bases de connaissance telles que DBpedia (AUER, BIZER, KOBILAROV, LEHMANN et IVES, 2007), Freebase (BOLLACKER, EVANS, PARITOSH, STURGE et TAYLOR, 2008) ou YAGO (SUCHANEK, KASNECI et WEIKUM, 2007), les connaissances factuelles sur les entités nommées sont devenues exploitables par des méthodes automatiques. Ceci est précisément ce sur quoi s'appuie notre contribution : la réconciliation temporelle.

7.3 Chaîne de traitement de la réconciliation temporelle

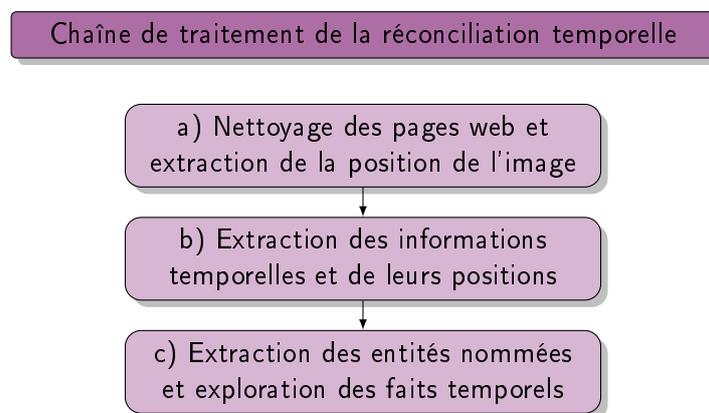


ILLUSTRATION 7.6: Diagramme présentant les étapes de la chaîne de traitement de notre approche.

Notre chaîne de traitement pour la réconciliation temporelle est décrite dans l'illustration 7.6. Elle comporte 3 étapes principales que nous détaillons ci-dessous :

a) **Nettoyage des pages web et extraction de la position de l'image :**

Nous avons nettoyé les pages web des scripts, des balises HTML, des styles, *etc.*⁴ Pour ce faire, nous avons utilisé une série d'expressions régulières et non un outil de l'état de l'art. Et ce afin de pouvoir parallèlement, détecter et stocker la position de l'image dans la page d'origine et dans la page nettoyée. Cette opération est nécessaire puisqu'elle permet d'apparier les pages nettoyées avec les annotations qui sont réalisées à partir des pages dans leurs affichages originaux.

4. Nous avons notamment converti les pages web en *UTF-8*.

b) Extraction des informations temporelles et de leurs positions :

Pour extraire du texte des informations temporelles explicites telles que des dates, décennies ou autres périodes de temps, nous avons choisi d'utiliser des expressions régulières. Nous avons conçu 3 expressions différentes spécifiques⁵ à notre tâche, dans lesquelles *A* représente un chiffre. Une première ayant la forme *AAAA* permettant de détecter des années *e.g.* 1919, 1976, *etc.* Une deuxième ayant la forme *AAA0s* permettant de détecter des décennies *e.g.* 1910s, 1970s, *etc.* Une troisième ayant la forme *AAAA-AAAA* permettant de détecter des périodes de temps *e.g.* 1919 – 1976, *etc.*

c) Extraction des entités nommées et exploration des faits temporels :

L'extraction de l'information temporelle sémantique est basée sur AIDA (HOFFART, YOSEF, BORDINO, FÜRSTENAU, PINKAL, SPANIOL, TANEVA, THATER et WEIKUM, 2011), qui désambiguïse les entités nommées provenant de la base de connaissances YAGO2 (HOFFART, SUCHANEK, BERBERICH et WEIKUM, 2010). Ces entités nommées nous permettent d'extraire des informations sémantiques en lien avec le contenu de nos pages web. En particulier, nous incorporons et nous utilisons, pour la réconciliation temporelle, les relations qui transportent l'information temporelle provenant des faits de *yagoLiteralFacts* (comme par exemple « *wasDestroyedOnDate* », « *wasBornOnDate* », *etc.*).

En fonction du type de l'entité nommée, et lorsque les données de temporalité sont présentes, nous pouvons avoir affaire, par exemple, aux dates de naissance et le cas échéant de décès d'une personne ou encore aux dates de début et de fin d'existence d'un produit, d'une société, d'un pays, d'un événement, *etc.* Dans le tableau 7.1, nous avons listé les marqueurs permettant d'inférer de tels intervalles de temps.

relations initiatrices	relations terminatrices
<i>startedOnDate</i>	<i>endedOnDate</i>
<i>startsExistingOnDate</i>	<i>endsExistingOnDate</i>
<i>wasBornOnDate</i>	<i>diedOnDate</i>
<i>wasCreatedOnDate</i>	<i>wasDestroyedOnDate</i>
<i>occursSince</i>	<i>occursUntil</i>

TABLEAU 7.1: Relations temporelles présentes dans *yagoLiteralFacts*. (source YAGO 2 (HOFFART, SUCHANEK, BERBERICH et WEIKUM, 2010))

Grâce à ces informations, nous disposons d'intervalles temporels. Selon la nature des données, ces intervalles peuvent être fermés, *i.e.* les dates de début et de fin sont connues, ou bien ouverts dans le cas contraire.

5. C'est pour cette raison que nous n'avons pas utilisé *Heideltime* (STRÖTGEN et GERTZ, 2010).

La dernière étape de notre chaîne de traitement est celle dans laquelle notre système « mime » le comportement humain. Dans cette phase, nous acceptons ou rejetons la prédiction en se basant sur sa conformité avec la réconciliation temporelle. La prédiction initiale peut être obtenue de plusieurs manières. Conformément aux énoncés précédents concernant les méthodes de déductions pragmatiques, la prédiction peut être la date la plus fréquente ou la date la plus proche de l'image. Dans le cas de la date la plus proche de l'image, cette proximité est calculée en nombre de caractères dans les pages web nettoyées, entre la position de l'information temporelle et la position de l'image « projetée ⁶ » dans la page nettoyée. Nous avons développé deux approches à partir de ces informations, l'une est stricte et l'autre souple.

7.3.1 Approche stricte

Nous faisons usage des « relations temporelles initiatrices » décrites dans le tableau 7.1, afin de déterminer *strictement* la borne inférieure d'une période temporelle d'acceptation et nous faisons respectivement la même chose avec les « relations temporelles terminatrices » pour déterminer sa borne supérieure. Dans ce cadre, nous considérons le temps comme une donnée continue et ordonnée. Parmi les entités nommées, nous observons toutes les relations initiatrices et nous choisissons la date la plus récente comme borne inférieure. À l'inverse, pour les relations terminatrices, nous choisissons la date la plus ancienne comme borne supérieure. Le fonctionnement de cette approche est similaire à celui décrit dans la frise de l'illustration 7.5 à la page 113.

À présent, observons une série d'exemples présentant le fonctionnement de notre approche en pratique. Dans toutes ces illustrations, les dates réelles, qui ont été fournies par les annotateurs ⁷ sont marquées en bleu ; tandis qu'en vert, apparaissent des histogrammes présentant le nombre d'occurrences de dates par année découverte par les expressions régulières. L'une d'entre elles est la date la plus proche de l'image. Elle est signalée par un trait rouge pointillé. Enfin, la zone rose hachurée indique la période temporelle qui est invalidée par le principe de réconciliation temporelle stricte.

Nous allons donc maintenant attirer l'attention du lecteur sur 3 illustrations représentant les 3 principaux comportements possibles :

- Un cas dans lequel nous acceptons la prédiction en nous basant sur la période de confiance donnée par la réconciliation temporelle (illustration 7.7) ;
- Un cas dans lequel nous rejetons la prédiction car elle se trouve hors de la période de confiance donnée par la réconciliation temporelle (illustration 7.8) ;

6. Lors de la suppression de la balise contenant l'image, sa position est conservée et transposée dans la page nettoyée.

7. Le processus d'annotation est détaillé dans le chapitre 5.

- Enfin, un cas dans lequel nous ne pouvons pas utiliser la réconciliation temporelle car les informations temporelles provenant des entités nommées sont discordantes (illustration 7.9).

Dans ce dernier cas, la borne supérieure obtenue à partir des informations temporelles est correcte puisqu'elle indique 1936 alors que la date réelle est 1920. En revanche la borne inférieure est incorrecte, puisqu'elle indique 1987, qui correspond à la relation temporelle initiatrice d'une des entités nommées détectées dans le texte. Ce fonctionnement constitue une limitation de cette méthode, et nous allons voir dans la sous-section 7.3.2 comment nous avons tenté de résoudre ce problème en relaxant le principe d'application de la réconciliation temporelle.

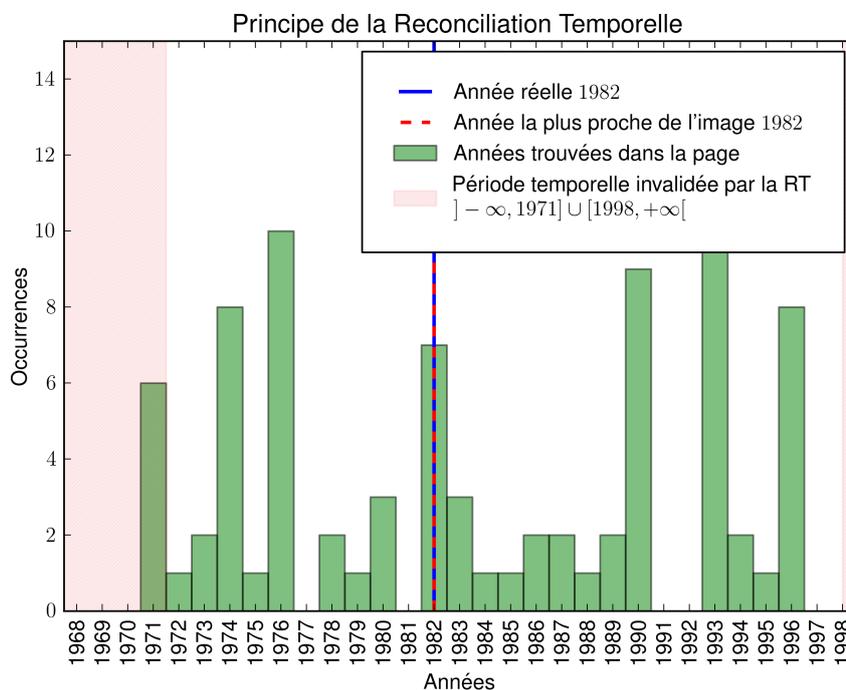


ILLUSTRATION 7.7: Exemple d'acceptation d'une prédiction par l'usage de la réconciliation temporelle stricte.

7.3.2 Approche souple

Nous avons développé une seconde approche permettant de déterminer une période temporelle de confiance. Cette méthode est une extension de l'approche stricte dont elle vise à améliorer la principale limitation : sa trop grande rigidité. Cette rigidité affecte le recoupement entre les différentes périodes qui, de ce fait, se retrouve souvent inexistant.

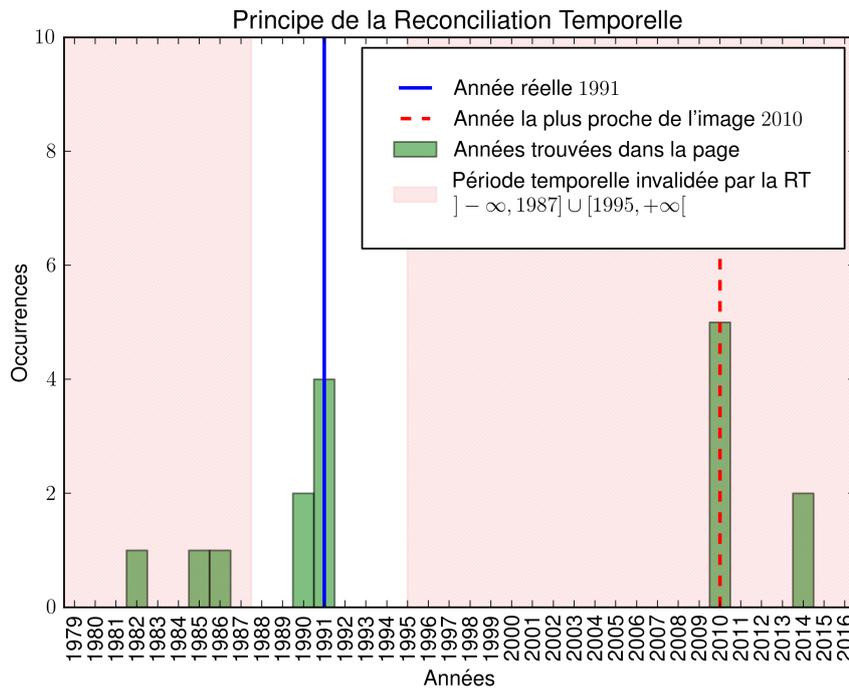


ILLUSTRATION 7.8: Exemple de rejet d'une prédiction par l'usage de la réconciliation temporelle stricte.

Dans cette approche, nous faisons usage des intervalles temporels associés aux entités nommées de manière à recouper les informations et à croiser les périodes dans le but de déterminer une densité de probabilité temporelle. L'idée d'un tel recoupement est inspirée de l'algèbre des intervalles temporels d'Allen (ALLEN, 1983).

Il arrive parfois que les intervalles temporels issus des entités nommées soient déconnectés entre eux, entraînant ainsi l'impossibilité de toute réconciliation temporelle (c.f. illustration 7.9). En réalité, dans ce genre de cas, des périodes temporelles communes entre les entités nommées sont bien souvent exploitables. Avec l'approche stricte il suffit d'une seule inconsistance temporelle pour empêcher le bon fonctionnement de la réconciliation. L'approche souple permet de contourner cette restriction.

Nous allons présenter notre approche de réconciliation temporelle souple en la comparant à la réconciliation temporelle stricte dans le cadre de 2 exemples concrets.

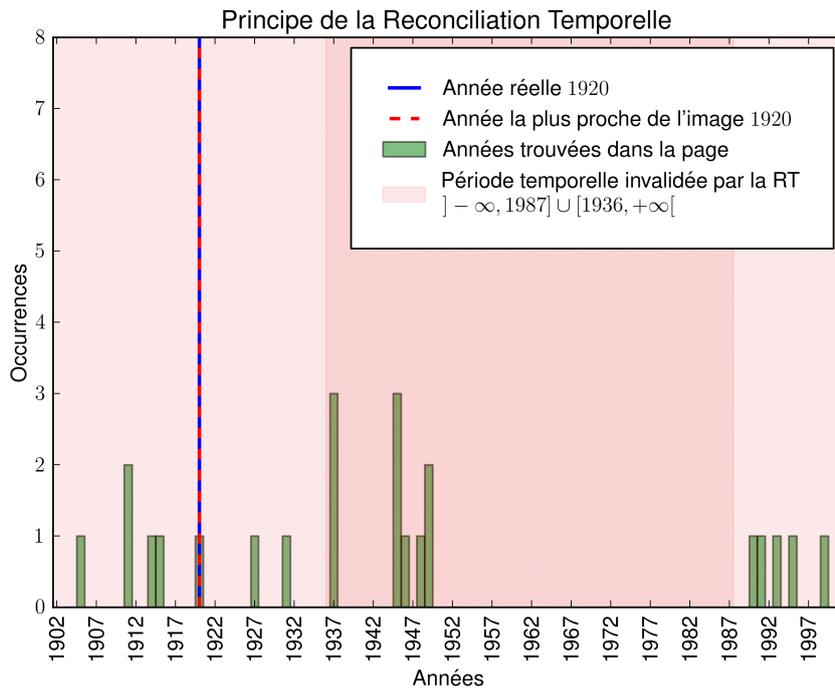


ILLUSTRATION 7.9: Exemple dans lequel la réconciliation temporelle stricte ne peut être appliquée du fait d'inconsistances temporelles.

Dans le premier exemple de l'illustration 7.10, ne_1 à ne_4 représentent les intervalles temporels déduits de 4 entités nommées. $RT_{stricte}$ représente la période temporelle validée par la réconciliation temporelle stricte. Dans cet exemple, il n'y a aucune période compatible, car les intervalles sont disjoints. En revanche, si l'on considère le nombre d'entités valides par unité de temps nous obtenons une densité de préférences temporelles, qui forme le cœur de notre méthode de réconciliation temporelle souple, notée ici RT_{souple} .

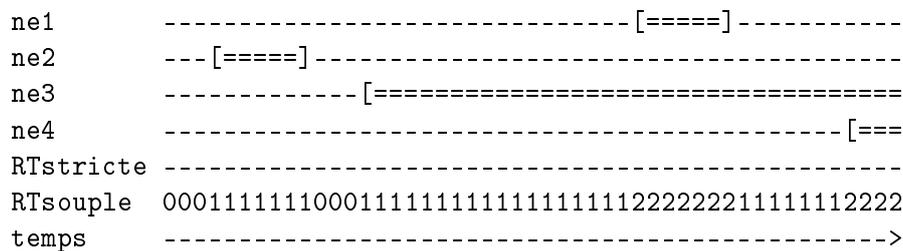


ILLUSTRATION 7.10: Réconciliation temporelle stricte vs. souple : cas des intervalles disjoints causés par des inconsistances temporelles entre les entités nommées.

Dans le second exemple de l'illustration 7.11, nous présentons un cas pour lequel la période d'acceptation résultante de la $RT_{stricte}$ est très courte. L'usage de la RT_{souple} permet alors de considérer d'autres dates proches du pic de concordance des intervalles.

```

ne1      ----- [=====] -----
ne2      --- [=====] -----
ne3      ----- [=====] -----
ne4      =====] -----
RTstricte -----
RTsouple 1112222233334443333222222222222222111111111111
temps    ----->

```

ILLUSTRATION 7.11: Réconciliation temporelle stricte vs. souple : cas des périodes d'intersection courtes.

Ainsi, l'approche souple fonctionne par comptage du nombre d'intervalles temporels en accord avec chaque période de temps. Le décompte ainsi produit est ensuite normalisé afin d'obtenir une densité de préférences temporelles, comparable à une densité de probabilité, qu'il est ensuite possible d'exploiter de manière plus riche que lors de la réconciliation temporelle stricte. Car tandis que l'approche stricte offre une validation binaire des dates candidates, les stratégies de l'approche souple permettent d'offrir une validation plus tolérante. Nous verrons dans la section 7.4 quelles sont les stratégies que nous avons employées pour exploiter cette densité de préférences temporelles.

7.4 Expériences et résultats

Nous avons mené des expériences sur la datation des photographies sur le jeu de données web-photo 1900-2000, que nous avons mis en place au cours de nos recherches. Ce jeu de données est détaillé dans la sous-section 5.2.3. Il est composé de 2 041 pages web composées de photographies prises au cours d'un ensemble de 100 années : [1900, 1999]. Dans les tableaux 7.2 et 7.3 nous reportons les résultats de nos expériences sur la réconciliation temporelle. Dans ces tableaux, « *mesure^M* » signifie que nous utilisons les versions *Macro* de ces mesures *i.e.* chaque classe est d'abord évaluée à part puis les résultats de toutes les classes sont moyennés entre eux. En plus des mesures d'évaluations classiques, décrites dans la section 4.3, nous indiquons le nombre de documents concernés : nombre de documents dont la date a été retrouvée (# bons), mal identifiée (# mauvais) et le nombre total de documents concernés par chaque approche (# total).

Nous avons évalué dans un premier temps, le cas des documents structurés en supposant que nous pouvons utiliser la date la plus proche de l'image comme prédiction initiale de notre processus de réconciliation temporelle. Puis dans un second temps, nous avons expérimenté le cas de documents non structurés dans lesquels il est impossible d'utiliser la proximité des informations temporelles à l'image.

méthodes \ mesures	ACC	MAE	MAE ^M	Précision ^M	Rappel ^M	Score F1 ^M	# bons	# mauvais	Total
Informations indiquant la qualité des données sources									
Date utilisée dans les requêtes	0.85	3	2,02	0.63	0.62	0.62	1 730	311	2 041
Scores de références obtenus par « déductions pragmatiques »									
Date la plus proche de l'image – Heuristique des moteurs de recherche	0.8	3,35	3,3	0.8	0.8	0.8	1 583	396	1 979
Date médiane	0.56	6,38	6,27	0.58	0.57	0.57	1 111	868	1 979
Date la plus fréquente	0.67	6,68	6,38	0.68	0.68	0.67	1 319	660	1 979
Date moyenne	0.34	7,59	7,54	0.38	0.35	0.35	674	1 305	1 979
Réconciliations temporelles à partir de la date la plus proche de l'image									
Réconciliation temporelle stricte	0.82	1,19	1,07	0.72	0.73	0.71	59	13	72
Réconciliation temporelle souple probabilité > 0.9	0.79	3,07	4,13	0.78	0.76	0.76	503	135	638
Réconciliation temporelle souple prédiction à partir du recoupement	0.88	2,57	2,73	0.88	0.88	0.87	760	106	866

TABLEAU 7.2: Évaluation de nos méthodes de réconciliation temporelle combinées à la date la plus proche de l'image dans le cas de documents structurés.

Documents structurés

Dans la partie supérieure du tableau nous pouvons observer les informations indiquant la qualité des données obtenues par les requêtes temporelles soumises à des moteurs de recherche. La qualité de ces données n'est pas optimale puisqu'un grand nombre d'informations quantitatives peuvent être utilisées par les moteurs de recherche pour retourner les résultats d'une requête temporelle, telles que des informations sur la résolution de l'image, la taille en mètres carrés d'un lieu, *etc.* Il peut aussi arriver que la date réelle ne soit pas mentionnée dans le texte de la page⁸.

À partir des données temporelles extraites des expressions régulières, nous avons déduit un certain nombre de dates qui nous ont servi de scores de références. Les performances de ces dates déduites sont très bonnes malgré leur déduction pragmatique. Nous les avons reportées dans la partie centrale du tableau 7.2. Nous pouvons observer que l'« heuristique des moteurs de recherche » que nous avons représentée comme la date la plus proche de l'image mentionnée dans le texte qui l'entoure, donne de très bons scores. Les résultats concernant la « Date médiane », la « Date la plus fréquente » et enfin la « Date moyenne » servent ici de repères témoignant de l'efficacité de la date la plus proche. En effet, les scores de ces heuristiques sont nettement inférieurs à ceux de la date la plus proche.

La partie basse du tableau 7.2 regroupe des résultats dans lesquels nous combinons notre méthode de réconciliation temporelle avec la date la plus proche de l'image puisque nous sommes en mesure de l'obtenir grâce à la structure du document.

La réconciliation temporelle stricte obtient une très bonne précision (82%) en offrant en outre une erreur moyenne (MAE^M) très faible : 1,07 années ; malheureusement cette approche n'est capable de proposer une date que dans 72 cas sur 2 041.

8. Sur notre jeu de données, la date réelle est présente dans 1 841 cas sur 2 041.

La réconciliation temporelle souple avec la date la plus proche est ici reportée dans les cas où la date la plus proche se trouve dans une période temporelle à forte probabilité, dans ce cas supérieur à 90%. Autrement dit, la date sera choisie, si et seulement si, elle appartient à une période recoupée par 90% des intervalles liés au document. Cette approche n'offre pas de nette amélioration, comparée aux résultats obtenus avec la seule date la plus proche de l'image. Cela est sûrement dû au bruit induit par l'ensemble des entités nommées combinées. En effet, un document peut regrouper un grand nombre d'entités nommées dans son contenu lesquelles peuvent appartenir à des périodes de temps distantes. Le résultat étant dans ce cas impacté par une trop grande volonté de correspondre à un maximum d'entités nommées sans distinction d'importance vis-à-vis de l'image concernée. Bien que de nature souple, cette approche reste trop stricte car trop sensible au bruit.

La prédiction à partir du recouvrement tire parti de la densité de probabilité construite par l'approche souple de la réconciliation temporelle. À partir de cette densité, le pic de probabilité est utilisé afin d'obtenir l'année la plus probable. Cette année n'étant pas systématiquement présente dans les pages web, nous choisissons l'année trouvée grâce aux expressions régulières la plus proche temporellement de celle-ci. Nous avons reporté les résultats obtenus avec cette approche dirigée par le recouvrement temporel lorsque l'année sélectionnée correspond de plus à la date la plus proche de l'image. Cette technique permet d'augmenter le nombre de documents pour lesquels la réconciliation temporelle peut s'appliquer tout en améliorant grandement les résultats. L'avantage de celle-ci face à la « réconciliation temporelle souple (probabilité > 0.9) » est que les cas pour lesquels la prédiction peut être accordée ne sont pas nécessairement les cas pour lesquels la majeure partie des entités nommées sont en accord ; ce qui offre une plus grande souplesse dans l'acceptation des prédictions face au bruit induit par l'ensemble des entités nommées.

Documents non structurés

méthodes \ mesures	<i>ACC</i>	<i>MAE</i>	<i>MAE^M</i>	<i>Précision^M</i>	<i>Rappel^M</i>	<i>Score F1^M</i>	# bons	# mauvais	total
Scores de références obtenus par « déductions pragmatiques »									
Date la plus fréquente	0.67	6,68	6,38	0.68	0.68	0.67	1 319	660	1 979
Réconciliations temporelles à partir de la date la plus fréquente									
Réconciliation temporelle stricte	0.66	3,48	2,7	0.6	0.57	0.56	52	27	79
Réconciliation temporelle souple probabilité > 0.9	0.67	6,13	6,78	0.67	0.64	0.63	419	206	625
Réconciliation temporelle souple prédiction à partir du recouvrement	0.89	2,63	2,89	0.89	0.88	0.88	728	93	821

TABLEAU 7.3: Évaluation de nos méthodes de réconciliation temporelle combinées à la date la plus fréquente dans le cas des documents non structurés.

Il existe de nombreux documents dans lesquels la structure de la page n'est pas accessible ou avec laquelle il ne sera pas possible d'établir un lien entre la position de l'image et la position des dates : cela peut être dû à la présence de scripts ou de techniques web avec lesquelles l'apparence de la page est calculée par le navigateur, *etc.* Si l'on se place dans ce contexte, alors l'« heuristique des moteurs de recherche », utilisant la date la plus proche, ne peut pas être exploitée. Et par conséquent, de ce fait, nous ne pouvons pas⁹ utiliser la date la plus proche de l'image comme prédiction initiale de notre réconciliation temporelle.

Dans de tels cas, il reste néanmoins possible d'employer la date la plus fréquente dans le document, mais cette technique ne donne pas de très bons résultats en comparaison de ceux obtenus lorsque la position de la date la plus proche de l'image est disponible (*c.f.* tableau 7.2).

C'est donc dans ces cas précis que la réconciliation temporelle permet une amélioration non négligeable. Nous avons vu dans le cas des documents structurés que la prédiction à partir du recoupement était la stratégie la plus efficace. Cette stratégie se vérifie une fois de plus, lorsque les prédictions sont dirigées par la probabilité maximale obtenue avec la réconciliation temporelle souple, la réconciliation permet d'être aussi efficace avec la date la plus fréquente qu'avec la date la plus proche de l'image. Le gain est particulièrement important si l'on considère l'*ACC* qui passe de 67% à 89%; ce gain est par ailleurs observable suivant toutes les mesures référencées dans le tableau 7.3. Le comportement de cette approche dans cette nouvelle configuration nous permet de vérifier l'intérêt de la méthode développée. De plus cela nous indique qu'il est important de maintenir une souplesse vis-à-vis du recoupement des intervalles temporels, cela permet d'éviter le bruit occasionné par les entités nommées.

7.5 Conclusion

Les approches d'horodatage que nous avons présentées « imitent » le comportement humain en exploitant l'information provenant des entités nommées. Comme nous l'avons signalé dans la section 7.4, nos résultats sont prometteurs et offrent la possibilité d'utiliser des informations sémantiques pour la datation. Cependant, cette réconciliation temporelle a un prix : alors que nous gagnons en précision, nous perdons en rappel. Les recherches futures devraient ainsi viser à accroître à la fois la précision et le rappel.

9. Pas de manière aisée.

Les principaux verrous sont, premièrement, la détection des entités nommées qui sont une source d'information déterminante pour la suite de notre approche. La détection des entités est victime de bruits venant parasiter le *continuum* temporel provenant par exemple d'entités récurrentes tels que des noms de villes dans notre jeu de données. Deuxièmement, la faible disponibilité des informations temporelles pour les entités nommées pose des problèmes pour le recoupement des informations temporelles. Il est donc délicat d'exploiter toutes les entités en raison du caractère incomplet des informations concernant la majorité d'entre elles. Troisièmement, l'exploitabilité des relations temporelles disponibles est aussi une limite, car, bien qu'il existe au total 5 types d'intervalles temporels distincts (c.f. tableau 7.1), dans nos expériences, les données temporelles n'ont concerné que 2 de ces intervalles (*wasBornOnDate/diedOnDate* et *wasCreatedOnDate/wasDestroyedOnDate*). Enfin, certaines relations peu fiables participent aux résultats finaux, ce qui nuit au bon fonctionnement de l'approche et suggère la mise en place d'une étape de sélection. Cette piste, que nous n'avons cependant pas eu l'opportunité de développer dans le cadre de ce travail doctoral, fait partie des nombreuses améliorations possibles pouvant se greffer autour de notre chaîne de traitement.

Conclusions et perspectives

Nous sommes à la genèse de cette problématique. La datation automatique de photographies partage des aspects à la croisée de domaines tels que l'histoire, les arts, les sciences humaines, l'informatique... La datation de photographies est donc un sujet nouveau, à large spectre qui intéresse notre société dans laquelle l'information est omniprésente et où la culture du passé suscite un engouement croissant. Le travail de mémoire sur les archives est motivé par la compréhension de notre histoire. La photographie tient une place majeure dans ce contexte. Il a donc été nécessaire dans notre étude de puiser dans des domaines très variés.

Nous avons compilé des informations historiques sur l'évolution de la photographie à travers le temps depuis ses origines. Cette étude contextualisante de l'évolution de la photographie et ses métadonnées nous a servi de base fondatrice dans nos hypothèses de travail. Nous espérons de plus que ces recherches pourront motiver de nouvelles études exploitant l'importance des caractéristiques historiques que nous avons mises en évidence.

Notre étude détaillée de l'état de l'art nous a permis d'inventorier les avancées concernant la datation de photographies. Ce sujet a la particularité d'avoir été peu étudié dans la vision par ordinateur comme dans le traitement des langues. Nous avons, par conséquent, parmi de nombreux axes de recherche possibles, suivi les pistes qui nous ont semblé les plus directes pour explorer ce sujet et apporter aux différentes communautés des contributions nouvelles et intéressantes.

Les avancées réalisées sur ce sujet sont principalement concentrées sur la définition de caractéristiques visuelles temporellement discriminantes pour les photographies. En complément de ces approches, nous avons développé un modèle permettant de tirer parti du caractère ordonné du temps et d'exploiter de manière adéquate ces descripteurs.

Compte tenu des observations que nous avons faites à propos des méthodes manuelles de datation de photographies nous avons pris connaissance du rôle prépondérant du texte dans la tâche de datation. Nous avons par conséquent entrepris d'investiguer cette vision nouvelle du problème au regard des travaux existants dans la communauté du traitement des langues. Pour ce faire, nous avons constitué un jeu de données visant à satisfaire ce besoin et combler cette lacune. Ce jeu permet la datation de photographies par les images, mais aussi, pour la première fois, par le texte les environnant.

Nous avons donc exploré cette nouvelle voie concernant la datation des photographies en proposant un modèle original permettant l'exploitation des textes à la manière d'un humain par des mécanismes de validation temporelle de périodes préférentielles.

Comme nous venons de le rappeler, la datation de photographies est une thématique large et peu explorée, alors les possibilités futures sont multiples. Nous allons à présent, illustrer dans de nouvelles perspectives de cette problématique quelques idées de développement se basant à la fois sur nos contributions en matière de datation mais aussi sur des travaux de l'état de l'art.

Datation par les objets présents dans l'image

Les méthodes actuelles de datation automatique de photographies (DIAS, MORENO, JATOWT et CAMPOS, 2012 ; PALERMO, HAYS et EFROS, 2012 ; FERNANDO, MUSELET, KHAN et TUYTELAARS, 2014 ; MARTIN, DOUCET et JURIE, 2014) reposent essentiellement sur l'analyse de l'aspect (par exemple par l'usage de caractéristiques visuelles fondées sur les couleurs) des photographies sans prendre en compte l'association des objets présents avec des périodes temporelles. Cependant, nous avons montré dans la section 3.1 que la datation manuelle reposait en grande partie sur l'analyse du contenu et des objets présents dans la photographie. À ce sujet, nous avons présenté dans la section 3.2 des travaux récents visant à analyser de façon automatique l'évolution des objets à travers le temps comme par exemple : les voitures (LEE, EFROS et HEBERT, 2013) ou les façades d'immeubles (LEE, MAISONNEUVE, CRANDALL, EFROS et SIVIC, 2015).

Il serait donc possible d'entrevoir la datation automatique de photographies par la reconnaissance d'objets caractéristiques d'une époque, s'inspirant ainsi du fonctionnement des méthodes manuelles. Le principal frein à de telles investigations reste la conception de bases de connaissances d'informations visuelles temporelles, difficulté que nous avons évoquée dans le chapitre 5. Ainsi, afin de déterminer un périmètre d'étude viable pour ces recherches, il sera nécessaire de limiter le champ des possibles concernant le contenu des photographies. En effet, tous les objets existants relèvent potentiellement d'une époque bien précise et seraient donc utiles pour la datation. Prendre l'exemple de la reconnaissance des vêtements serait donc un premier pas dans cette direction. La littérature de la datation manuelle de photographies sur ce sujet est fournie (PEACOCK, 2008 ; *Victoria and Albert Museum* ; *English Museums Association*). Parallèlement, certains travaux de vision par ordinateur permettent déjà d'identifier le type de vêtement (BORRÀS, TOUS, LLADÓS et VANRELL, 2003). Il serait souhaitable de poursuivre des recherches dans cette direction afin de développer des caractéristiques sensibles aux configurations vestimentaires qui ont marqué les années (voir l'illustration 7.12) basées par exemple sur l'extraction de *patches* de textures associant motifs récurrents et couleurs caractéristiques vestimentaires, sur les coupes des vêtements, *etc.*



ILLUSTRATION 7.12: Évolution des tenues de bain féminines entre 1900 et 1919. (source *Met Museum* et *Abiti Antichi collection*)

Néanmoins ce domaine reste très diversifié car les vêtements sont dépendants du genre, du style, de l'âge, de l'activité, de la localisation, de la culture, *etc.* Toutes ces déclinaisons rendront sans doute la tâche très compliquée. Une autre piste possible dans cette même direction pourrait être par exemple la reconnaissance de logotypes temporels dans les photographies. La littérature sur la temporalité des logotypes est à première vue accessible (voir l'illustration 7.13).

L'évolution temporelle des logotypes est un sujet qui passionne certains internautes qui produisent des frises chronologiques imagées (voir l'illustration 7.13) qui pourraient servir de ressources informatives pour la constitution de bases d'apprentissage.



ILLUSTRATION 7.13: Résultats de la requête « logo evolution ». (source Google images)

La reconnaissance de logotypes a déjà été étudiée dans la communauté de la vision par ordinateur (RUSIÑOL et LLADÓS, 2009), cette tâche rentre dans le cadre de la reconnaissance de symboles (LLADÓS, MARTÍ et VILLANUEVA, 2001 ; LLADÓS, VALVENY, SÁNCHEZ et MARTÍ, 2002). L'extension de telles méthodes pourrait donc être envisagée et plusieurs usages se dessineraient alors. L'objectif pourrait être la reconnaissance dans des photographies (voir l'illustration 7.15) des différents logotypes utilisés à travers le temps par les marques. À partir de ces informations, plusieurs schémas se dessinent. Premièrement, nous pourrions exploiter les informations temporelles extraites de ceux-ci. Les logotypes reconnus, pourraient être associés à des intervalles temporels et fonctionner ainsi à l'égal de la combinaison des intervalles temporels et des périodes d'acceptation telles que déterminées par notre processus de réconciliation temporelle (présenté dans le chapitre 7). Deuxièmement, les logotypes pourraient être regroupés par marque d'appartenance, par exemple tous les logos représentant la marque BP (voir l'illustration 7.14).

STATION SERVICE



ILLUSTRATION 7.14: Évolution entre 1900 et 2000 de logotypes associés à 3 stations services répandues : BP, Shell et Texaco. (source longue-vue.com)

Ainsi à partir des informations de chaque marque, nous pourrions identifier les entités nommées représentatives de ces marques. Ces entités nommées nous permettraient d'obtenir des informations temporelles élevant le contenu textuel à un niveau sémantique, de façon similaire au travail que nous avons présenté dans le chapitre 7.



ILLUSTRATION 7.15: Piccadilly Circus à Londres en 1962, nous apercevons de nombreux logotypes dont celui de BP qui correspond au modèle qui a été utilisé entre 1958 et 1989. (photographie issue de notre jeu de données web-photo 1900-2000)

Amélioration de la réconciliation temporelle

Nous avons mis en place dans le chapitre 5 un jeu de données issu du web, nommé web-photo 1900-2000, permettant d'évaluer les méthodes de datation de photographies par l'usage du texte et de l'image. Nous avons proposé dans le chapitre 7 une première approche permettant la datation de photographies provenant de notre jeu de données. Notre approche permet la validation de la datation par un processus de réconciliation temporelle utilisant des informations de temporalité extraites de multiples entités nommées, issues des textes.

Notre méthode est robuste et nous envisageons dès à présent, compte tenu des résultats de nos expériences reportés dans la section 7.4 de l'améliorer encore.

Lorsque notre système accepte toutes les entités nommées, il devient difficile d'obtenir une réponse¹⁰. De même, lorsque notre système tente de considérer un maximum d'entités nommées, le gain de cette approche n'est pas à la hauteur de nos attentes¹¹. Ces stratégies représentent de nouvelles approches intéressantes pour exploiter le caractère temporel des entités nommées. Cependant l'effet de bord de telles stratégies provient de la présence d'entités « parasites » venant déstabiliser le *continuum* temporel et ainsi gêner le fonctionnement de la réconciliation temporelle.

Concevoir une méthode de sélection des entités nommées participant à la réconciliation temporelle, permettrait d'éliminer le bruit induit par les entités nommées « parasites ».

Une première approche empirique pourrait être de sélectionner les entités en fonction de la distance qui les séparent de l'image que nous souhaitons dater. Ainsi une telle hypothèse suppose que le contenu textuel proche de l'image est le plus informatif pour déduire la date de prise de vue de celle-ci. Nous avons montré dans nos expériences que les informations temporelles proches de l'image sont en effet pertinentes. Il serait intéressant de savoir si les informations sémantiques se comportent de la même manière. En se basant sur de tels objectifs, d'autres améliorations pourraient en plus considérer le nombre d'occurrences des entités au sein à la fois de la page, mais aussi du corpus dans son ensemble. L'objectif d'une telle approche serait d'adapter les techniques associant la fréquence des termes dans les documents à leur fréquence dans la collection, avec une ambition similaire à celle de la mesure *tf-idf* (RAMOS, 1999).

Une seconde approche, plus sophistiquée, et inspirée de travaux récents sur les entités nommées viserait à sélectionner les entités les plus pertinentes en fonction de leur évolution temporelle. En effet, les travaux de BOUVIER et BELLOT (2015) ont récemment proposé une méthode capable de suivre l'évolution temporelle des entités nommées. Cette méthode leur permet ainsi de sélectionner les documents traitant de nouveautés. D'une part, l'usage de ces méthodes avec le fonctionnement inverse serait sans doute capable de proposer un filtrage des entités récentes. D'autre part, une extension de ces méthodes permettrait aussi de suivre l'évolution des entités nommées dans le passé par exemple en s'inspirant des observations faites par WEIKUM, NTARMOS, SPANIOL, TRIANTAFILLOU, BENCZÚR, KIRKPATRICK, RIGAUX et WILLIAMSON (2011). Ainsi, en fonction d'une première période de préférence temporelle donnée par la réconciliation temporelle, de telles techniques permettraient d'élire les entités nommées les plus représentatives de la période et donc de mettre les entités parasites à l'écart.

Afin de déterminer l'importance des entités nommées, une troisième approche est de s'appuyer sur le type de l'entité considérée. En effet, les entités en lien avec des personnes ou des bâtiments sont sans doute à considérer de manière distincte, quant à leur rôle dans le processus de réconciliation temporelle.

10. Référence à l'approche stricte, décrite dans la sous-section 7.3.1.

11. Référence à l'approche souple, décrite dans la sous-section 7.3.2.

Élire des entités nommées préférentielles permettrait sans doute d'améliorer les approches de réconciliation temporelle. Cette élection peut être effectuée de diverses manières, en s'appuyant par exemple sur la distance qui les séparent des images, leur nombre d'occurrences, *etc.*

Si nous sommes capables de déterminer l'importance des entités nommées dans chaque document, nous pouvons alors tenter de répondre au problème du manque d'information temporelle qui a lieu dans certains cas. À ce stade, nous pouvons envisager d'étendre sémantiquement les entités les plus importantes avec d'autres entités ayant un fort lien sémantique avec elles, et n'étant pas présentes dans le document concerné. Les entités nommées ainsi rassemblées permettraient de compléter le spectre temporel validé par le processus de réconciliation temporelle.

Réflexions et pistes pour la datation visuelle

Nous pouvons citer d'autres techniques auxquelles nous avons pensé au cours de cette thèse pour obtenir des indices temporels utiles à la datation de photographies.

Comme nous l'avons indiqué dans la section 6.5 à la page 102, FERNANDO, MUSELET, KHAN et TUYTELAARS (2014) ont proposé de nouveaux descripteurs qui tentent d'exploiter des caractéristiques discriminantes du type d'appareil ayant réalisé le cliché. Cependant, les performances de leurs descripteurs n'ont pas été vérifiées sur un jeu de données de classification d'appareil photographique. Selon nos observations un tel jeu n'existe pas dans la communauté de la vision par ordinateur. C'est pourquoi, cette tâche n'a encore jamais été traitée. Constituer un jeu de données basé sur des modèles d'appareils les plus utilisés dans l'histoire de la photographie et tester l'efficacité de ces descripteurs seraient capital. S'il s'avère que ces descripteurs bas niveau, ou d'autres descripteurs de plus haut niveau de l'état de l'art, peuvent avoir une haute performance sur cette nouvelle tâche, ceci pourrait grandement contribuer aux méthodes de datation de photographies. En effet, comme nous l'avons indiqué dans la sous-section 3.1.1 à la page 32, reconnaître avec une haute fidélité le type d'appareil ou la technologie photographique ayant réalisé un cliché, offrira de solides indices temporels quant à la période de réalisation de celui-ci. Si nous pouvons détecter le modèle ayant réalisé le cliché nous pourrions obtenir des informations temporelles via des entités nommées et les associer à un système de bases de connaissances temporelles tel que nous l'avons exploité dans le chapitre 7.

Dans un autre registre THOMAS et KOVASHKA (2015) ont comparé des descripteurs bas et haut niveau dans une nouvelle tâche. Cette tâche consiste à reconnaître le photographe ayant réalisé la photographie. Les auteurs ont proposé un jeu de 100 000 photographies pour lesquelles les photographes auteurs sont connus. Supposons que cette technique puisse être étendue de manière semi-supervisée afin de reconnaître et découvrir de nouveaux photographes non encore répertoriés dans ce catalogue restreint. Cette technique pourrait, similairement à notre méthode, permettre d'obtenir des informations sur la personne ayant réalisé le cliché telles que sa date de naissance, la période d'activité en temps que photographe, le cas échéant la date de son décès, *etc.* Ces informations serviraient alors à alimenter la période de confiance mise en place par la réconciliation temporelle que nous avons introduite dans le chapitre 7.

Dans cette thèse nous nous sommes donc concentrés sur la recherche et la classification d'information temporelle, plus précisément la datation de photographies. Cette thématique est relativement nouvelle et finalement peu explorée. Il nous a été possible de proposer des contributions qui permettront de cadrer les recherches futures. En effet, ce sujet reste ouvert et de nombreux travaux peuvent se greffer sur nos observations et nos contributions.

Les principaux points qui seront importants de considérer peuvent en outre concerner l'élaboration de :

- modèles de modélisation des problèmes liés à la temporalité (*e.g.* grâce à de nouvelles méthodes de classification ordinale) ;
- descripteurs plus sensibles à la temporalité des photographies ;
 - visuellement (*e.g.* descripteurs couleurs, patch en niveau de gris...),
 - textuellement (*e.g.* entités nommées temporelles, vocabulaires temporels...).
- mesures d'évaluation adaptées à l'usage de séparations temporelles non régulières.

– Paul MARTIN.

Liste des illustrations

1.1	Exemple d'image ancienne insérée dans une page web : l'objectif est de prédire automatiquement la date de prise de vue de cette image en se basant à la fois sur le contenu visuel de l'image et sur le texte qui l'entourne. (source <i>lavanguardia.com</i>)	6
1.2	Opposition entre contenu et aspect d'une photographie selon les époques.	7
1.3	Les meilleurs résultats de recherche sont principalement focalisés sur la ou les dimensions thématiques prédominantes. (source <i>Google images</i> août 2015)	13
1.4	Les particularités géographiques et culturelles sont déduites et contraintes par la langue de la requête. (source <i>Google images</i> août 2015)	14
1.5	Exemple de résultats de recherche pour la requête « <i>New York</i> ».	16
1.6	La continuité temporelle d'un sujet, ici une personnalité : Michael Jackson.	17
1.7	Exemple d'application possible, l'évolution temporelle d'un sujet, ici une personnalité : Michael Jackson [1958, 2009].	17
1.8	Exemple d'application possible, la réorganisation temporelle d'un album photographique numérisé. Ici $t_1 < t_2 < t_3 < t_4$ sont des marqueurs temporels non identifiés mais ordonnés.	18
2.1	Publicité de 1889 pour le <i>Kodak No.1</i> créé en 1888. (source <i>Kodak</i>)	22
2.2	Le recto d'un <i>ferrotype</i> digitalisé dans son ensemble par un appareil photographique numérique. (source <i>flickr</i> « <i>The Commons</i> »)	24
2.3	Publicité de 1915 pour le film <i>Kodak Autographique</i> . (source <i>Kodak</i>)	25
2.4	Exemples de différents horodatages et formats de dates selon les technologies photographiques.	26
2.5	Ajout d'horodatages et de géolocalisation sur les images avec une application pour smartphone. (source « <i>TimeStamp Photo with GPS Address Location</i> » sur <i>itunes</i>)	27
2.6	Exemple de données EXIF capturées avec un appareil photographique numérique.	28
3.1	L'évolution des technologies photographiques. (source <i>popchartlab.com</i>)	32

3.2	Différents filigranes se trouvant au verso d'impressions photographiques, variant selon les époques photographiques et les constructeurs. (source des images WEAVER et LONG (2009))	37
4.1	Formule OC_{β}^{γ} . (source CARDOSO et SOUSA (2011))	63
5.1	Quelques institutions participant au projet « <i>The Commons</i> ».	76
5.2	Une frise chronologique d'images de rues de notre base.	80
5.3	24 photographies de rues appartenant à notre jeu de données de documents, web-photo 1900-2000.	85
5.4	Distribution du nombre d'images par année dans notre jeu de données web-photo 1900-2000.	86
5.5	Distribution du nombre d'images par décennie dans notre jeu de données web-photo 1900-2000.	87
5.6	Distribution du nombre d'images en fonction de leurs formats dans notre jeu de données web-photo 1900-2000.	87
5.7	Distribution du nombre d'images en fonction de leurs résolutions dans notre jeu de données web-photo 1900-2000.	88
6.1	Chaîne de traitement en apprentissage statistique supervisé : l'exemple de la classification d'images.	90
6.2	Exemples d'images de la collection de PALERMO, HAYS et EFROS (2012).	100
7.1	Dans cet exemple aucune mention de date n'apparaît, seule la légende peut donner une information de temporalité : « Cérémonie d'ouverture des Jeux Olympiques d'été de Los Angeles ». (source summergames.ap.org)	108
7.2	Dans cet exemple plusieurs mentions de dates apparaissent : 1984, 10/08/1984 et 1972. (source olympic.org)	110
7.3	Photographie d'illustration d'un article du 5 octobre 2012 sur le lancement de l' <i>iPhone</i> . (source darkroom.baltimoresun.com)	111
7.4	Informations extraites des entités nommées détectées. (source <i>Wikipédia</i>)	112
7.5	Démonstration conceptuelle de l'approche de la réconciliation temporelle.	113
7.6	Diagramme présentant les étapes de la chaîne de traitement de notre approche.	114
7.7	Exemple d'acceptation d'une prédiction par l'usage de la réconciliation temporelle stricte.	117
7.8	Exemple de rejet d'une prédiction par l'usage de la réconciliation temporelle stricte.	118
7.9	Exemple dans lequel la réconciliation temporelle stricte ne peut être appliquée du fait d'inconsistances temporelles.	119
7.10	Réconciliation temporelle stricte vs. souple : cas des intervalles disjoints causés par des inconsistances temporelles entre les entités nommées.	119

7.11 Réconciliation temporelle stricte vs. souple : cas des périodes d'intersection courtes.	120
7.12 Évolution des tenues de bain féminines entre 1900 et 1919. (source <i>Met Museum</i> et <i>Abiti Antichi collection</i>)	127
7.13 Résultats de la requête « logo evolution ». (source <i>Google images</i>)	128
7.14 Évolution entre 1900 et 2000 de logotypes associés à 3 stations services répandues : <i>BP</i> , <i>Shell</i> et <i>Texaco</i> . (source <i>longue-vue.com</i>)	128
7.15 Piccadilly Circus à Londres en 1962, nous apercevons de nombreux logotypes dont celui de <i>BP</i> qui correspond au modèle qui a été utilisé entre 1958 et 1989. (photographie issue de notre jeu de données <i>web-photo 1900-2000</i>)	129
A.1 Frise chronologique géologique. (source <i>talkorigins.org</i>)	157
A.2 Première photographie « Point de vue du Gras », depuis la fenêtre de la maison de Joseph Nicéphore Niépce. (support ayant imprimé la lumière). (source <i>Harry Ransom Humanities Research Center, The University of Texas, Austin</i>)	158
A.3 Premier selfie pris par l'américain Robert Cornelius [1809, 1893] en 1839.	159
A.4 Première photographie couleur prise par Thomas Sutton en 1861.	159
A.5 Première photographie digitale (scannée), cette image représente le fils de Russell Kirsch, l'inventeur du scanner à plat en 1957.	160
B.1 Le déclin de la pellicule. (source <i>technologyreview.com</i>)	163
B.2 La chute du chiffre d'affaires des impressions avec l'arrivée de la photographie en numérique. (source <i>technologyreview.com</i>)	164
B.3 <i>Kodak</i> dépose le bilan en 2012. (source <i>reuters.com</i>)	165
D.1 Images datées dans l'intervalle [1899, 2016] selon les données EXIF extraites à partir des 150 000 images recueillies lors de la conception de notre jeu de données web, voir chapitre 5.	172
D.2 Images datées dans l'intervalle [1851, 2016] selon les données XMP extraites à partir des 150 000 images recueillies lors de la conception de notre jeu de données web, voir chapitre 5.	174
F.1 Interface web de « National Portrait Gallery ».	181
F.2 Une photographie de la base de la bibliothèque de Toulouse.	182
F.3 Détails disponibles pour une photographie de la base de la bibliothèque de Toulouse.	182
F.4 Présentation base Gallica, de nombreux supports différents recueils et albums	183
F.5 Présentation base Gallica (suite) la notice complète d'un album de photographies	184
F.6 Présentation de la base de la Médiathèque.	185
F.7 Images de la base <i>flickr</i> « <i>The Commons</i> ».	186

F.8	Capture d'écran du site du journal Ouest-France représentant une rue de Caen photographiée le 31 décembre 1925 ainsi que son texte descriptif associé.	187
G.1	Image <i>4152965237.jpg</i> associée au fichier XML <i>4152965237.xml</i> . (source flickr « <i>The Commons</i> »)	189
G.2	Contenu du fichier XML <i>4152965237.xml</i> associé à l'image <i>4152965237.jpg</i> . (source flickr « <i>The Commons</i> »)	190
G.3	Présentation de l'interface d'annotation réalisée.	191
G.4	Répartition par année de la collection complète des 42 310 documents numériques que nous avons extraits.	192
G.5	Répartition par année de la collection des 8 831 documents numériques que nous avons extraits pour lesquels la précision est à l'année près.	193
G.6	Répartition par année de la collection 1 093 documents numériques que nous avons extraits et que nous avons annotés comme représentatifs de photographies de rues.	193
H.1	Résultats de recherche avec le moteur de recherche <i>Google images</i> pour la requête Tokyo 1993.	197
H.2	Début de l'article du blog. (source http://reggies-words.blogspot.com/)	198
H.3	Fin de l'article du blog. (source http://reggies-words.blogspot.com/) .	199
H.4	Exemple de fichier <i>snippet</i> retourné par <i>Google image</i> que nous avons stocké dans des fichiers json. Le contenu du fichier « <i>0_5945de39-75fe-4570-8f23-2d7a18d5276a.json</i> ».	200
H.5	Image récupérée associée au contenu du blog.	200
H.6	Distribution du nombre d'images en fonction de leurs formats dans notre jeu de données <i>web-photo</i> 1900-2000 pour la décennie 1900.	201
H.7	Distribution du nombre d'images en fonction de leurs formats dans notre jeu de données <i>web-photo</i> 1900-2000 pour la décennie 1910.	202
H.8	Distribution du nombre d'images en fonction de leurs formats dans notre jeu de données <i>web-photo</i> 1900-2000 pour la décennie 1920.	202
H.9	Distribution du nombre d'images en fonction de leurs formats dans notre jeu de données <i>web-photo</i> 1900-2000 pour la décennie 1930.	203
H.10	Distribution du nombre d'images en fonction de leurs formats dans notre jeu de données <i>web-photo</i> 1900-2000 pour la décennie 1940.	203
H.11	Distribution du nombre d'images en fonction de leurs formats dans notre jeu de données <i>web-photo</i> 1900-2000 pour la décennie 1950.	204
H.12	Distribution du nombre d'images en fonction de leurs formats dans notre jeu de données <i>web-photo</i> 1900-2000 pour la décennie 1960.	204
H.13	Distribution du nombre d'images en fonction de leurs formats dans notre jeu de données <i>web-photo</i> 1900-2000 pour la décennie 1970.	205

H.14	Distribution du nombre d'images en fonction de leurs formats dans notre jeu de données web-photo 1900-2000 pour la décennie 1980.	205
H.15	Distribution du nombre d'images en fonction de leurs formats dans notre jeu de données web-photo 1900-2000 pour la décennie 1990.	206
I.1	Matrice de confusion de l'article de (FERNANDO, MUSELET, KHAN et TUY-TELAARS, 2014).	208

Liste des tableaux

1.1	Opposition entre vocabulaire et style textuel selon les époques.	8
4.1	Différences de conversion entre indices, scores, étiquettes et rangs (avec K classes et n instances).	59
5.1	Statistiques du jeu de données.	84
6.1	Ré-étiquetage des exemples d'apprentissage des sous-classifieurs avec la stratégie <i>Avant contre Après</i>	94
6.2	Impact du ré-étiquetage sur la balance des exemples d'apprentissage des sous-classifieurs <i>Avant contre Après</i>	94
6.3	Calculs des poids appliqués aux classes ré-étiquetées des sous-classifieurs obtenus avec la stratégie <i>Avant contre Après</i>	95
6.4	Pondérations des exemples d'apprentissage par le poids $w(n, k)$ de l'équation 6.1.	96
6.5	Performance de l'approche proposée sur le jeu de données de PALERMO, HAYS et EFROS (2012), ainsi que les comparaisons avec les approches connexes.	102
6.6	Résultats de FERNANDO, MUSELET, KHAN et TUYTELAARS (2014) que nous avons calculés à partir de la matrice de confusion reportée dans leur article.	104
7.1	Relations temporelles présentes dans <i>yagoLiteralFacts</i> . (source YAGO 2 (HOFFART, SUCHANEK, BERBERICH et WEIKUM, 2010))	115
7.2	Évaluation de nos méthodes de réconciliation temporelle combinées à la date la plus proche de l'image dans le cas de documents structurés.	121
7.3	Évaluation de nos méthodes de réconciliation temporelle combinées à la date la plus fréquente dans le cas des documents non structurés.	122
A.1	Frise chronologique retraçant l'évolution de la photographie [1826, 1888].	158
A.2	Frise chronologique retraçant l'évolution de la photographie [1888, 1957].	160
A.3	Frise chronologique retraçant l'évolution de la photographie [1957, 2000].	161

D.1	Données EXIF hors années [1899, 2016] extraites à partir des 150 000 images recueillies lors de la conception de notre jeu de données web, voir chapitre 5.	173
D.2	Données XMP hors années [1851, 2016] extraites à partir des 150 000 images recueillies lors de la conception de notre jeu de données web, voir chapitre 5.	175
H.1	Liste des 50 villes au monde les plus visitées au monde pendant l'année 2013. (source MasterCard)	196
J.1	Comparatif des résultats de notre méthode de classification ordinale présentée dans le chapitre 6 avec ceux de SÁNCHEZ-MONEDERO, GUTIÉRREZ, ANTONIO, TINO et HERVÁS-MARTÍNEZ (2013) en terme de $ACC_{\text{écart type}}$. . .	210
J.2	Comparatif des résultats de notre méthode de classification ordinale présentée dans le chapitre 6 avec ceux de SÁNCHEZ-MONEDERO, GUTIÉRREZ, ANTONIO, TINO et HERVÁS-MARTÍNEZ (2013) en terme de $MAE_{\text{écart type}}$. . .	210
J.3	Comparatif des résultats de notre méthode de classification ordinale présentée dans le chapitre 6 avec ceux de SÁNCHEZ-MONEDERO, GUTIÉRREZ, ANTONIO, TINO et HERVÁS-MARTÍNEZ (2013) en terme de $AMAE_{\text{écart type}}$. . .	211
J.4	Comparatif des résultats de notre méthode de classification ordinale présentée dans le chapitre 6 avec ceux de SÁNCHEZ-MONEDERO, GUTIÉRREZ, ANTONIO, TINO et HERVÁS-MARTÍNEZ (2013) en terme de $\tau_b_{\text{écart type}}$. . .	211

Références bibliographiques

ALLEN, James F. (nov. 1983). « Maintaining Knowledge About Temporal Intervals ». Dans : *Commun. ACM* 26.11, p. 832–843. ISSN : 0001-0782. DOI : 10.1145/182.358434. URL : <http://doi.acm.org/10.1145/182.358434>.

ANGLIN, Paul M. et Ramazan GENÇAY (1996). « Semiparametric estimation of a hedonic price function ». Dans : *Journal of Applied Econometrics* 11.6, p. 633–648. ISSN : 1099-1255. DOI : 10.1002/(SICI)1099-1255(199611)11:6<633::AID-JAE414>3.0.CO;2-T. URL : [http://dx.doi.org/10.1002/\(SICI\)1099-1255\(199611\)11:6<633::AID-JAE414>3.0.CO;2-T](http://dx.doi.org/10.1002/(SICI)1099-1255(199611)11:6<633::AID-JAE414>3.0.CO;2-T).

ARDIZZONE, Edoardo, H DINDO et G MAZZOLA (2009). *Content-Based Image Retrieval as Validation for Defect Detection in Old Photos*. T. 2009.

ARIE, Ben David. *Weka Ordinal, real-world datasets (ERA, ESL, LEV and SWD)*. URL : <http://www.cs.waikato.ac.nz/~ml/weka/datasets.html>.

AUER, Sören, Christian BIZER, Georgi KOBILAROV, Jens LEHMANN et Zachary IVES (2007). « DBpedia : A Nucleus for a Web of Open Data ». Dans : *In 6th International Semantic Web Conference, Busan, Korea*. Springer, p. 11–15.

BACCIANELLA, Stefano. *TripAdvisor Users review (15 763) dataset*. URL : <http://patty.isti.cnr.it/~baccianella/reviewdata/index.php?download>.

BACCIANELLA, Stefano, Andrea ESULI et Fabrizio SEBASTIANI (2009a). « Evaluation measures for ordinal regression ». Dans : *Intelligent Systems Design and Applications, 2009. ISDA'09. Ninth International Conference on*. IEEE, p. 283–287.

BACCIANELLA, Stefano, Andrea ESULI et Fabrizio SEBASTIANI (2009b). « Multi-facet Rating of Product Reviews ». Dans : *Proceedings of the 31th European Conference on IR Research on Advances in Information Retrieval*. ECIR '09. Toulouse, France : Springer-Verlag, p. 461–472. ISBN : 978-3-642-00957-0. DOI : 10.1007/978-3-642-00958-7_41. URL : http://dx.doi.org/10.1007/978-3-642-00958-7_41.

BACHE, K. et M. LICHMAN (2013). *UCI Machine Learning Repository*. URL : <http://archive.ics.uci.edu/ml>.

BAEZA-YATES, Ricardo, Berthier RIBEIRO-NETO et al. (1999). *Modern information retrieval*. T. 463. ACM press New York.

BENAVENTE, Robert, Maria VANRELL et Ramon BALDRICH (2008). « Parametric fuzzy sets for automatic color naming ». Dans : *J. Opt. Soc. Am. A* 25.10, p. 2582–2593. DOI : 10.1364/JOSAA.25.002582. URL : <http://josaa.osa.org/abstract.cfm?URI=josaa-25-10-2582>.

BERBERICH, K., S.J BEDATHUR, O. ALONSO et G. WEIKUM (2010). « A Language Modeling Approach for Temporal Information Needs ». Dans : *32nd European Conference on Information Retrieval*, p. 13–25.

BERENGUER ARROYO, Eloi et Robert Benavente i VIDAL (2013). « Aplicació per a la predicció automàtica de la data d'adquisició d'una imatge ». Mém.de maîtr. Universitat Autònoma de Barcelona. Escola d'Enginyeria.

Bibliothèque de Toulouse – Photographies anciennes. Base accessible ici : <http://rosalis.bibliotheque.toulouse.fr/index.php?pages/photographiesanciennes>. Dernier accès août 2015.

Bibliothèque numérique Gallica. Base accessible ici : <http://gallica.bnf.fr>. Dernier accès août 2015.

BISHOP, Christopher M (2006). *Pattern recognition and machine learning*. springer.

BLITZER, John, Mark DREDZE et Fernando PEREIRA (2007). « Biographies, Bollywood, Boom-boxes and Blenders : Domain Adaptation for Sentiment Classification ». Dans : *Association for Computational Linguistics*. Prague, Czech Republic.

BOLLACKER, Kurt, Colin EVANS, Praveen PARITOSH, Tim STURGE et Jamie TAYLOR (2008). « Freebase : A Collaboratively Created Graph Database for Structuring Human Knowledge ». Dans : *Proceedings of the 2008 ACM SIGMOD International Conference on Management of Data*. SIGMOD '08. Vancouver, Canada : ACM, p. 1247–1250. ISBN : 978-1-60558-102-6. DOI : 10.1145/1376616.1376746. URL : <http://doi.acm.org/10.1145/1376616.1376746>.

BORRÀS, Agnès, Francesc TOUS, Josep LLADÓS et María VANRELL (2003). « High-Level Clothes Description Based on Colour-Texture and Structural Features. » Dans : *IBPRIA*. Sous la dir. de Francisco J. Perales LÓPEZ, Aurélio C. CAMPILHO, Nicolas Pérez de la BLANCA et Alberto SANFELIU. T. 2652. Lecture Notes in Computer Science. Springer, p. 108–116. ISBN : 3-540-40217-9. URL : <http://dblp.uni-trier.de/db/conf/ibpria/ibpria2003.html#BorrasTLV03>.

BOUVIER, Vincent et Patrice BELLOT (2015). « Modèles de langue adaptatifs et méta-critères pour le filtrage de documents et le suivi temporel d'entités ». Dans : *Document Numérique* 18.1, p. 75–96. DOI : 10.3166/dn.18.1.75-96. URL : <http://dx.doi.org/10.3166/dn.18.1.75-96>.

BOYNTON, Robert M. et Conrad X. OLSON (1990). « Saliency of chromatic basic color terms confirmed by three measures ». Dans : *Vision Research* 30.9, p. 1311–1317. ISSN : 0042-6989. DOI : [http://dx.doi.org/10.1016/0042-6989\(90\)90005-6](http://dx.doi.org/10.1016/0042-6989(90)90005-6). URL : <http://www.sciencedirect.com/science/article/pii/0042698990900056>.

BRILL, Eric David (1993). « A Corpus-based Approach to Language Learning ». UMI Order No. GAX93-31757. Thèse de doct. Philadelphia, PA, USA.

BROWN, Peter F., John COCKE, Stephen A. Della PIETRA, Vincent J. Della PIETRA, Fredrick JELINEK, John D. LAFFERTY, Robert L. MERCER et Paul S. ROOSSIN (juin 1990). « A Statistical Approach to Machine Translation ». Dans : *Comput. Linguist.* 16.2, p. 79–85. ISSN : 0891-2017. URL : <http://dl.acm.org/citation.cfm?id=92858.92860>.

CARDOSO, Jaime S. et Maria J. CARDOSO (juin 2007). « Towards an intelligent medical system for the aesthetic evaluation of breast cancer conservative treatment ». Dans : *Artif. Intell. Med.* 40.2, p. 115–126. ISSN : 0933-3657. DOI : 10.1016/j.artmed.2007.02.007. URL : <http://dx.doi.org/10.1016/j.artmed.2007.02.007>.

CARDOSO, Jaime S. et Joaquim F. Pinto da COSTA (déc. 2007). « Learning to Classify Ordinal Data : The Data Replication Method ». Dans : *Journal of Machine Learning Research* 8, p. 1393–1429. ISSN : 1532-4435. URL : <http://dl.acm.org/citation.cfm?id=1314498.1314546>.

CARDOSO, Jaime S et Ricardo SOUSA (2011). « Measuring the performance of ordinal classification ». Dans : *International Journal of Pattern Recognition and Artificial Intelligence* 25.08, p. 1173–1195.

CHANG, Chih-Chung et Chih-Jen LIN (2011). « LIBSVM : A library for support vector machines ». Dans : *ACM Transactions on Intelligent Systems and Technology* 2 (3). Software available at <http://www.csie.ntu.edu.tw/~cjlin/libsvm>, 27 :1–27 :27.

CHANG, Kuang-Yu, Chu-Song CHEN et Yi-Ping HUNG (2011). « Ordinal hyperplanes ranker with cost sensitivities for age estimation ». Dans : *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, p. 585–592.

CHATZICHRISTOFIS, Savvas A. et Yiannis S. BOUTALIS (2008a). « CEDD : Color and Edge Directivity Descriptor : A Compact Descriptor for Image Indexing and Retrieval ». Dans : *Proceedings of the 6th International Conference on Computer Vision Systems. ICVS'08*. Santorini, Greece : Springer-Verlag, p. 312–322. ISBN : 3-540-79546-4, 978-3-540-79546-9. URL : <http://dl.acm.org/citation.cfm?id=1788524.1788559>.

CHATZICHRISTOFIS, Savvas A. et Yiannis S. BOUTALIS (2008b). « FCTH : Fuzzy Color and Texture Histogram - A Low Level Feature for Accurate Image Retrieval ». Dans : *Proceedings of the 2008 Ninth International Workshop on Image Analysis for Multimedia Interactive Services. WIAMIS '08*. Washington, DC, USA : IEEE Computer Society, p. 191–196. ISBN : 978-0-7695-3130-4. DOI : 10.1109/WIAMIS.2008.24. URL : <http://dx.doi.org/10.1109/WIAMIS.2008.24>.

CHEN, Wei, Tie yan LIU, Yanyan LAN, Zhi ming MA et Hang LI (2009). « Ranking Measures and Loss Functions in Learning to Rank ». Dans : *Advances in Neural Information Processing Systems 22*. Sous la dir. d'Y. BENGIO, D. SCHUURMANS, J. LAFFERTY, C. WILLIAMS et A. CULOTTA, p. 315–323. URL : http://books.nips.cc/papers/files/nips22/NIPS2009_0493.pdf.

CHEN, Xiang-Rong et HongJiang ZHANG (2003). « Photo time-stamp detection and recognition ». Dans : *7th International Conference on Document Analysis and Recognition (ICDAR 2003), 2-Volume Set, 3-6 August 2003, Edinburgh, Scotland, UK*, p. 319–322. DOI : 10.1109/ICDAR.2003.1227681. URL : <http://dx.doi.org/10.1109/ICDAR.2003.1227681>.

CHENG, Jianlin, Zheng WANG et G. POLLASTRI (2008). « A neural network approach to ordinal regression ». Dans : *Neural Networks, 2008. IJCNN 2008. (IEEE World Congress on Computational Intelligence). IEEE International Joint Conference on*, p. 1279–1284. DOI : 10.1109/IJCNN.2008.4633963.

CHU, Wei et Zoubin GHAMRANI (2004). « Gaussian processes for ordinal regression ». Dans :

CHU, Wei et S Sathiya KEERTHI (2005). « New approaches to support vector ordinal regression ». Dans : *Proceedings of the 22nd international conference on Machine learning*. ACM, p. 145–152.

CIEPLINSKI, Leszek (2001). « MPEG-7 Color Descriptors and Their Applications. » Dans : CAIP. Sous la dir. de Wladyslaw SKARBK. T. 2124. Lecture Notes in Computer Science. Springer, p. 11–20. ISBN : 3-540-42513-6. URL : <http://dblp.uni-trier.de/db/conf/caip/caip2001.html#Cieplinski01>.

COE, B. et M. HAWORTH-BOOTH (1983). *A Guide to Early Photographic Processes*. Victoria & Albert Museum, London.

COSTA, Joaquim F. Pinto da, Hugo ALONSO et Jaime S. CARDOSO (2008). « The unimodal model for the classification of ordinal data ». Dans : *Neural Networks* 21.1, p. 78–91. ISSN : 0893-6080. DOI : 10.1016/j.neunet.2007.10.003. URL : <http://www.sciencedirect.com/science/article/pii/S089360800700202X>.

COSTA, Joaquim F. Pinto da, Ricardo SOUSA et Jaime S. CARDOSO (2010). « An all-at-once Unimodal SVM Approach for Ordinal Classification ». Dans : *Machine Learning and Applications (ICMLA), 2010 Ninth International Conference on*. IEEE, p. 59–64.

CRAMMER, Koby et Yoram SINGER (2001). « Pranking with Ranking ». Dans : *Advances in Neural Information Processing Systems 14*. MIT Press, p. 641–647.

DIAS, Gaël, José G. MORENO, Adam JATOWT et Ricardo CAMPOS (2012). « Temporal Web Image Retrieval ». Dans : *String Processing and Information Retrieval - 19th International Symposium, SPIRE 2012, Cartagena de Indias, Colombia, October 21-25, 2012. Proceedings*, p. 199–204. DOI : 10.1007/978-3-642-34109-0_21. URL : http://dx.doi.org/10.1007/978-3-642-34109-0_21.

DIAS, Gaël Harry, Mohammed HASANUZZAMAN, Stéphane FERRARI et Yann MATHET (2014). « TempoWordNet for Sentence Time Tagging ». Dans : *Proceedings of the Companion Publication of the 23rd International Conference on World Wide Web Companion*. WWW Companion '14. Seoul, Korea : International World Wide Web Conferences Steering Committee, p. 833–838. ISBN : 978-1-4503-2745-9. DOI : 10.1145/2567948.2579042. URL : <http://dx.doi.org/10.1145/2567948.2579042>.

DIY Photo Dating Wizard. <http://www.cartedevisite.co.uk/DatingWizard/DatingWizard-Introduction.php>. Dernier accès août 2015.

DOERSCH, Carl, Saurabh SINGH, Abhinav GUPTA, Josef SIVIC et Alexei A. EFROS (2012). « What Makes Paris Look like Paris ? » Dans : *ACM Transactions on Graphics (SIGGRAPH)* 31.4, 101 :1–101 :9.

DOLSAK, Bojan et Stephen MUGGLETON (1992). « The Application of Inductive Logic Programming to Finite Element Mesh Design ». Dans : *Inductive Logic Programming*. Academic Press, p. 453–472.

DUIVESTELJN, Wouter et Ad FEELDERS (2008). « Nearest Neighbour Classification with Monotonicity Constraints ». Dans : *Proceedings of the 2008 European Conference on Machine Learning and Knowledge Discovery in Databases - Part I. ECML PKDD '08*. Antwerp, Belgium : Springer-Verlag, p. 301–316. ISBN : 978-3-540-87478-2. DOI : 10.1007/978-3-540-87479-9_38. URL : http://dx.doi.org/10.1007/978-3-540-87479-9_38.

DZEROSKI, Saso, Hendrik BLOCKEEL, Boris KOMPARE, Stefan KRAMER, Bernhard PFAHRINGER et Wim Van LAER (1999). « Experiments in Predicting Biodegradability ». Dans : *Applied Artificial Intelligence*. Springer, p. 80–91.

ELKAN, Charles (2001). « The Foundations of Cost-sensitive Learning ». Dans : *17th International Joint Conference on Artificial Intelligence. IJCAI*. Seattle, WA, USA : Morgan Kaufmann Publishers Inc., p. 973–978. ISBN : 1-55860-812-5, 978-1-558-60812-2. URL : <http://dl.acm.org/citation.cfm?id=1642194.1642224>.

English Museums Association. <http://www.museumsassociation.org/>. Dernier accès août 2015.

FERNANDO, Basura, Damien MUSELET, Rahat KHAN et Tinne TUYTELAARS (2014). « Color features for dating historical color images ». Dans : *2014 IEEE International Conference on Image Processing, ICIP 2014, Paris, France, October 27-30, 2014*, p. 2589–2593. DOI : 10.1109/ICIP.2014.7025524. URL : <http://dx.doi.org/10.1109/ICIP.2014.7025524>.

FERNANDO, Basura, Tatiana TOMMASI et Tinne TUYTELAARS (2015). « Location recognition over large time lags ». Dans : *Computer Vision and Image Understanding*, p. –. ISSN : 1077-3142. DOI : <http://dx.doi.org/10.1016/j.cviu.2015.05.016>. URL : <http://www.sciencedirect.com/science/article/pii/S107731421500137X>.

FRANK, Eibe et Mark HALL (2001). « A simple approach to ordinal classification ». Dans : *12th European Conference on Machine Learning (ECML)*. Springer, p. 145–156.

GAMMETER, Stephan, Lukas BOSSARD, Till QUACK et Luc J. Van GOOL (2009). « I know what you did last summer : object-level auto-annotation of holiday snaps. » Dans : *ICCV. IEEE*, p. 614–621. URL : <http://dblp.uni-trier.de/db/conf/iccv/iccv2009.html#GammeterBQG09>.

GAUDETTE, Lisa et Nathalie JAPKOWICZ (2009). « Evaluation methods for ordinal classification ». Dans : *Advances in Artificial Intelligence*, p. 207–210.

GENG, Xin, Zhi-Hua ZHOU et Kate SMITH-MILES (déc. 2007). « Automatic Age Estimation Based on Facial Aging Patterns ». Dans : *IEEE Trans. Pattern Anal. Mach. Intell.* 29.12, p. 2234–2240. ISSN : 0162-8828. DOI : 10.1109/TPAMI.2007.70733. URL : <http://dx.doi.org/10.1109/TPAMI.2007.70733>.

GROSS, Oskar, Antoine DOUCET et Hannu TOIVONEN (mar. 2013). « Named Entity Filtering based on Concept Association Graphs ». Dans : *14th International Conference in Computational Linguistics and Intelligent Text Processing (CICLing 2013)*. Sous la dir. d'Alexander GELBUKH. Samos, Greece, 11 pages. URL : <https://hal.archives-ouvertes.fr/hal-01073640>.

GUY, Stephen J. (2008). « Automatic Timestamp Detection and Removal from Digital Photos ». Computational Photography Final Report - <https://wwwx.cs.unc.edu/~sjguy/CompPhoto/>.

HALL, Mark, Eibe FRANK, Geoffrey HOLMES, Bernhard PFAHRINGER, Peter REUTEMANN et Ian H. WITTEN (nov. 2009). « The WEKA data mining software : an update ». Dans : *SIGKDD Explor. Newsl.* 11.1, p. 10–18. ISSN : 1931-0145. DOI : 10.1145/1656274.1656278. URL : <http://doi.acm.org/10.1145/1656274.1656278>.

« Handbook of environmental degradation rates » (1991). Dans :

HARDING, Colin. *Blog du « National Media Museum » de Grande Bretagne*. <http://blog.nationalmediamuseum.org.uk/tag/date-a-photograph/>. Dernier accès août 2015.

HARRINGTON, Edward F. (2003). « Online Ranking/Collaborative Filtering Using the Perceptron Algorithm ». Dans : *Proceedings of the 20th International Conference on Machine Learning*, p. 250–257.

HAYS, James et Alexei A. EFROS (2008). « im2gps : estimating geographic information from a single image ». Dans : *Proceedings of the IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*.

HE, Sheng, P. SAMMARA, J. BURGERS et L. SCHOMAKER (2014). « Towards Style-Based Dating of Historical Documents ». Dans : *Frontiers in Handwriting Recognition (ICFHR), 2014 14th International Conference on*, p. 265–270. DOI : 10.1109/ICFHR.2014.52.

HERBRICH, Ralf, Thore GRAEPEL et Klaus OBERMAYER (1999). « Large margin rank boundaries for ordinal regression ». Dans : *Advances in Neural Information Processing Systems*, p. 115–132.

HERBRICH, Ralf, Thore GRAEPEL, Klaus OBERMAYER et al. (1999). *Regression models for ordinal data : A machine learning approach*. Citeseer.

HOFFART, Johannes, Fabian M. SUCHANEK, Klaus BERBERICH et Gerhard WEIKUM (nov. 2010). *YAGO2 : a spatially and temporally enhanced knowledge base from Wikipedia*. Research Report MPI-I-2010-5-007. Stuhlsatzenhausweg 85, 66123 Saarbrücken, Germany : Max-Planck-Institut für Informatik.

HOFFART, Johannes, Mohamed Amir YOSEF, Ilaria BORDINO, Hagen FÜRSTENAU, Manfred PINKAL, Marc SPANIOL, Bilyana TANEVA, Stefan THATER et Gerhard WEIKUM (2011). « Robust Disambiguation of Named Entities in Text ». Dans : *Proceedings of the Conference on Empirical Methods in Natural Language Processing*. EMNLP '11. Edinburgh, United Kingdom : Association for Computational Linguistics, p. 782–792. ISBN : 978-1-937284-11-4. URL : <http://dl.acm.org/citation.cfm?id=2145432.2145521>.

JATOWT, Adam, Ching-Man AU YEUNG et Katsumi TANAKA (2013). « Estimating Document Focus Time ». Dans : *Proceedings of the 22Nd ACM International Conference on Conference on Information & Knowledge Management*. CIKM '13. San Francisco, California, USA : ACM, p. 2273–2278. ISBN : 978-1-4503-2263-8. DOI : 10.1145/2505515.2505655. URL : <http://doi.acm.org/10.1145/2505515.2505655>.

JINDAL, Nitin et Bing LIU (2008). « Opinion spam and analysis ». Dans : *Proceedings of the international conference on Web search and web data mining*. ACM, p. 219–230.

JOLY, Guy. *Datation des anciennes photographies*. http://guy.joly1.free.fr/datation_photos.html. Dernier accès août 2015.

JONG, prof.dr. F.M.G. de, H. RODE et D. HIEMSTRA (2005). « Temporal Language Models for the Disclosure of Historical Text ». Dans : *Humanities, computers and cultural heritage : Proceedings of the XVIth International Conference of the Association for History and Computing (AHC 2005)*. Imported from EWI/DB PMS [db-utwente :inpr :0000003683]. Amsterdam, The Netherlands : Royal Netherlands Academy of Arts et Sciences, p. 161–168. URL : <http://doc.utwente.nl/66448/>.

KANHABUA, Nattiya et Kjetil NØRVÅG (2008). « Improving Temporal Language Models for Determining Time of Non-timestamped Documents ». Dans : *Research and Advanced Technology for Digital Libraries*. Sous la dir. de Birte CHRISTENSEN-DALSGAARD, Donatella CASTELLI, Bolette AMMITZBØLL JURIK et Joan LIPPINCOTT. T. 5173. Lecture Notes in Computer Science. Springer Berlin Heidelberg, p. 358–370. ISBN : 978-3-540-87598-7. DOI : 10.1007/978-3-540-87599-4_37. URL : http://dx.doi.org/10.1007/978-3-540-87599-4_37.

KENDALL, M. G. (1938). « A New Measure of Rank Correlation ». Dans : *Biometrika* 30.1/2, p. 81–93. ISSN : 00063444. URL : <http://www.jstor.org/stable/2332226>.

KNERR, S., L. PERSONNAZ et G. DREYFUS (1990). « Single-layer learning revisited : a stepwise procedure for building and training a neural network ». English. Dans : *Neurocomputing*. Sous la dir. de Françoise Fogelman SOULIÉ et Jeanny HÉRAULT. T. 68. NATO ASI Series. Springer Berlin Heidelberg, p. 41–50. ISBN : 978-3-642-76155-3. DOI : 10.1007/978-3-642-76153-9_5. URL : http://dx.doi.org/10.1007/978-3-642-76153-9_5.

KRAMER, Stefan, Gerhard WIDMER, Bernhard PFAHRINGER et Michael de GROEVE (2001). « Prediction of ordinal classes using regression trees ». Dans : *Fundamenta Informaticae* 47.1-2, p. 1–13.

KRAPAC, J., M. ALLAN, J. VERBEEK et F. JURIE (2010). « Improving Web-Image Search Results Using Query-Relative Classifiers ». Dans : *23rd Conference on Computer Vision & Pattern Recognition*, p. 1094–1101.

LEE, Stefan, Nicolas MAISONNEUVE, David GRANDALL, Alexei EFROS et Josef SIVIC (2015). « Linking past to present : Discovering style in two centuries of architecture ». Dans : *IEEE International Conference on Computational Photography (ICCP)*.

LEE, Yong Jae, Alexei A. EFROS et Martial HEBERT (2013). « Style-Aware Mid-level Representation for Discovering Visual Connections in Space and Time ». Dans : *IEEE International Conference on Computer Vision, ICCV 2013, Sydney, Australia, December 1-8, 2013*, p. 1857–1864. DOI : 10.1109/ICCV.2013.233. URL : <http://dx.doi.org/10.1109/ICCV.2013.233>.

LEJEUNE, Gaël (oct. 2013). « Veille épidémiologique multilingue : une approche parcimonieuse au grain caractère fondée sur le genre textuel ». Theses. Université de Caen. URL : <https://hal.archives-ouvertes.fr/tel-01074940>.

LEJEUNE, Gaël, Romain BRIXTEL et Charlotte LECLUZE (2015). « Évaluation intrinsèque et extrinsèque du nettoyage de pages Web ». Dans : *Actes de la 22e conférence sur le Traitement Automatique des Langues Naturelles (TALN'2015)*, p. 411–417.

LERMAN, Israël-César (1992). « Conception et analyse de la forme limite d'une famille de coefficients statistiques d'association entre variables relationnelles. II ». fr. Dans : *Mathématiques et Sciences Humaines* 119, p. 75–100. URL : <http://eudml.org/doc/94431>.

LI, Ling et H LIN (2007). « Ordinal regression by extended binary classification ». Dans : *Advances in neural information processing systems* 19, p. 865.

LIN, H.T., Y.S. ABU-MOSTAFA, California Institute of Technology. Division of ENGINEERING et Applied SCIENCE (2008). *From Ordinal Ranking to Binary Classification*. CIT theses. California Institute of Technology. URL : <http://books.google.fr/books?id=6uYqQwAACAAJ>.

LIN, Stephen, Jinwei GU, Shuntaro YAMAZAKI et Heung-Yeung SHUM (25 oct. 2004). « Radiometric Calibration from a Single Image. » Dans : *CVPR (2)*, p. 938–945. URL : <http://dblp.uni-trier.de/db/conf/cvpr/cvpr2004-2.html#LinGYS04>.

LLADÓS, Josep, Enric MARTÍ et Juan José VILLANUEVA (2001). « Symbol recognition by error-tolerant subgraph matching between region adjacency graphs ». Dans : *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 23.10, p. 1137–1143.

LLADÓS, Josep, Ernest VALVENY, Gemma SÁNCHEZ et Enric MARTÍ (2002). « Symbol recognition : Current advances and perspectives ». Dans : *Graphics Recognition Algorithms and Applications*. Springer Berlin Heidelberg, p. 104–128.

LOCHBAUM, Karen E. et Lynn A. STREETER (nov. 1989). « Comparing and Combining the Effectiveness of Latent Semantic Indexing and the Ordinary Vector Space Model for Information Retrieval ». Dans : *Inf. Process. Manage.* 25.6, p. 665–676. ISSN : 0306-4573. DOI : 10.1016/0306-4573(89)90100-3. URL : [http://dx.doi.org/10.1016/0306-4573\(89\)90100-3](http://dx.doi.org/10.1016/0306-4573(89)90100-3).

LOUPY, Claude de, Patrice BELLOT, Marc EL-BÈZE et Pierre-Francois MARTEAU (1998). « Query Expansion and Classification of Retrieved Documents ». Dans : *Proceedings of The Seventh Text REtrieval Conference, TREC 1998, Gaithersburg, Maryland, USA, November 9-11, 1998*, p. 382–389.

MANNING, Christopher D. et Hinrich SCHÜTZE (1999). *Foundations of Statistical Natural Language Processing*. Cambridge, MA, USA : MIT Press. ISBN : 0-262-13360-1.

MARTIN, Paul (2013). « Datation d'images pour la recherche temporelle d'information ». Dans : *CORIA*. Sous la dir. de Catherine BERRUT. UNINE, p. 353–358.

MARTIN, Paul, Antoine DOUCET et Frédéric JURIE (2014). « Dating Color Images with Ordinal Classification ». Dans : *Proceedings of International Conference on Multimedia Retrieval*. ACM, p. 447.

MARTIN, Paul, Marc SPANIOL et Antoine DOUCET (2015). « Temporal Reconciliation for Dating Photographs Using Entity Information ». Dans : *prochainement : CIKM - Workshop on Exploiting Semantic Annotations in Information Retrieval (ESAIR) Melbourne AUSTRALIA. 2015.*

Médiathèque de l'Architecture et du Patrimoine – Base Mémoire : Archives photographiques. Base accessible ici : <http://www.mediathèque-patrimoine.culture.gouv.fr/>. Dernier accès août 2015.

MIGNON, Alexis (2015). *python-flickr-api*. Un wrapper python orienté objet à l'API de *flickr* accessible ici : <https://github.com/alexis-mignon/python-flickr-api>.

MILLER, George A. (nov. 1995). « WordNet : A Lexical Database for English ». Dans : *Commun. ACM* 38.11, p. 39–41. ISSN : 0001-0782. DOI : 10.1145/219717.219748. URL : <http://doi.acm.org/10.1145/219717.219748>.

MYERS, Jerome L, Arnold D WELL et Robert Frederick LORCH (2010). *Research design and statistical analysis*. Routledge.

NADEAU, David et Satoshi SEKINE (2007). « A survey of named entity recognition and classification ». Dans : *Linguisticae Investigationes* 30.1. Publisher : John Benjamins Publishing Company, p. 3–26. URL : <http://www.ingentaconnect.com/content/jbp/li/2007/00000030/00000001/art00002>.

National Portrait Gallery. Base accessible ici : <http://npgportraits.si.edu/emuseumCAP/code/emuseum.asp>. Dernier accès août 2015.

NAVIGLI, Roberto (fév. 2009). « Word Sense Disambiguation : A Survey ». Dans : *ACM Comput. Surv.* 41.2, 10 :1–10 :69. ISSN : 0360-0300. DOI : 10.1145/1459352.1459355. URL : <http://doi.acm.org/10.1145/1459352.1459355>.

NITIN, Jindal et Liu BING. *Amazon.com Users review (5.6 million) dataset*. URL : <http://liu.cs.uic.edu/download/data/>.

OLIVA, Aude et Antonio TORRALBA (mai 2001). « Modeling the Shape of the Scene : A Holistic Representation of the Spatial Envelope ». Dans : *Int. J. Comput. Vision* 42.3, p. 145–175. ISSN : 0920-5691. URL : <http://portal.acm.org/citation.cfm?id=598462>.

PALERMO, Frank, James HAYS et Alexei A. EFROS (2012). « Dating Historical Color Images ». Dans : *ECCV* (6), p. 499–512.

- PEACOCK, J. (2008). *La mode de 1900 à nos jours*. Thames & Hudson.
- PLATT, J. (2000). « Probabilistic outputs for support vector machines and comparison to regularized likelihood methods ». Dans : *Advances in Large Margin Classifiers*.
- POTHARST, Rob et JanC. BIOCH (1999). « A Decision Tree Algorithm for Ordinal Classification ». English. Dans : *Advances in Intelligent Data Analysis*. Sous la dir. de DavidJ. HAND, JoostN. KOK et MichaelR. BERTHOLD. T. 1642. Lecture Notes in Computer Science. Springer Berlin Heidelberg, p. 187–198. ISBN : 978-3-540-66332-4. DOI : 10.1007/3-540-48412-4_16. URL : http://dx.doi.org/10.1007/3-540-48412-4_16.
- PUSTEJOVSKY, James, José CASTAÑO, Robert INGRIA, Roser SAURÍ, Robert GAIZAUSKAS, Andrea SETZER et Graham KATZ (2003). « Timeml : Robust specification of event and temporal expressions in text ». Dans : *in Fifth International Workshop on Computational Semantics (IWCS-5)*.
- RAMOS, Juan (1999). *Using TF-IDF to Determine Word Relevance in Document Queries*.
- REMATAS, Konstantinos, Basura FERNANDO, Tatiana TOMMASI et Tinne TUYTELAARS (2013). *Does Evolution cause a Domain Shift ?*
- RITTER, Alan, Sam CLARK, Oren ETZIONI et al. (2011). « Named entity recognition in tweets : an experimental study ». Dans : *Proceedings of the Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics, p. 1524–1534.
- RUSIÑOL, Marçal et Josep LLADÓS (2009). « Logo Spotting by a Bag-of-words Approach for Document Categorization. » Dans : *ICDAR*. IEEE Computer Society, p. 111–115. ISBN : 978-0-7695-3725-2. URL : <http://dblp.uni-trier.de/db/conf/icdar/icdar2009.html#Rusinoll09>.
- SAGOT, Benoît, Karën FORT, Gilles ADDA, Joseph MARIANI et Bernard LANG (juin 2011). « Un turc mécanique pour les ressources linguistiques : critique de la myriadisation du travail parcellisé ». Dans : *TALN'2011 - Traitement Automatique des Langues Naturelles*. Montpellier, France. URL : <https://hal.inria.fr/inria-00617067>.
- SÁNCHEZ-MONEDERO, Javier, GUTIÉRREZ, Pedro ANTONIO, Peter TINO et César HERVÁS-MARTÍNEZ (2013). « Exploitation of Pairwise Class Distances for Ordinal Classification ». Dans : *Neural Computation* 25.9, MIT Press.
- SCALISE, A. et A.C. BLOSE (2010). *Dating images from scanned watermarks*. US Patent App. 12/491,268. URL : <http://www.google.com.ar/patents/US20100329575>.

SCALISE, A. et A.C. BLOSE (2012). *Dating images from scanned watermarks*. US Patent 8,290,205. URL : <http://www.google.com.ar/patents/US8290205>.

SCHMID, Helmut (1994). « Probabilistic part-of-speech tagging using decision trees ». Dans : *Proceedings of the international conference on new methods in language processing*. T. 12. Citeseer, p. 44–49.

SHASHUA, Amnon et Anat LEVIN (2003). *Ranking with Large Margin Principle : Two Approaches*.

SIEGLER, Robert S (1976). « Three aspects of cognitive development ». Dans : *Cognitive Psychology* 8.4, p. 481–520.

SINGH, Dinesh, Phillip G FEBBO, Kenneth ROSS, Donald G JACKSON, Judith MANOLA, Christine LADD, Pablo TAMAYO, Andrew A RENSHAW, Anthony V D'AMICO, Jerome P RICHIE et al. (2002). « Gene expression correlates of clinical prostate cancer behavior ». Dans : *Cancer cell* 1.2, p. 203–209.

SPEARMAN, C. (1904). « The Proof and Measurement of Association Between Two Things ». Dans : *American Journal of Psychology* 15, p. 88–103.

STRÖTGEN, Jannik, Omar ALONSO et Michael GERTZ (2012). « Identification of Top Relevant Temporal Expressions in Documents ». Dans : *Proceedings of the 2Nd Temporal Web Analytics Workshop*. TempWeb '12. Lyon, France : ACM, p. 33–40. ISBN : 978-1-4503-1188-5. DOI : 10.1145/2169095.2169102. URL : <http://doi.acm.org/10.1145/2169095.2169102>.

STRÖTGEN, Jannik et Michael GERTZ (2010). « HeidelTime : High Quality Rule-based Extraction and Normalization of Temporal Expressions ». Dans : *Proceedings of the 5th International Workshop on Semantic Evaluation*. SemEval '10. Los Angeles, California : Association for Computational Linguistics, p. 321–324. URL : <http://dl.acm.org/citation.cfm?id=1859664.1859735>.

STURGES, JULIA et T.W.ALLAN WHITFIELD (1997). « Salient Features of Munsell Colour Space as a Function of Monolexic Naming and Response Latencies ». Dans : *Vision Research* 37.3, p. 307–313. ISSN : 0042-6989. DOI : [http://dx.doi.org/10.1016/S0042-6989\(96\)00170-8](http://dx.doi.org/10.1016/S0042-6989(96)00170-8). URL : <http://www.sciencedirect.com/science/article/pii/S0042698996001708>.

SUCHANEK, Fabian M., Gjergji KASNECI et Gerhard WEIKUM (2007). « Yago : A Core of Semantic Knowledge ». Dans : *Proceedings of WWW*, p. 697–706.

flickr « The Commons ». Base accessible ici : <http://www.flickr.com/commons/>. Dernier accès août 2015.

THE MUSEUM OF MODERN ART. *Modern Photographs 1909–1949 : The Thomas Walther Collection*. <http://www.moma.org/objectphoto>. Dernier accès août 2015. « *The Thomas Walther Collection–341 photographs by 148 artists–represents the innovative vision of the 1920s and ‘30s, a transformative period of modern photography and the foundation of our photo-based world.* »

THOMAS, Christopher et Adriana KOVASHKA (2015). « Who’s Behind the Camera ? Identifying the Authorship of a Photograph ». Dans : *arXiv preprint arXiv :1508.05038*.

TORRALBA, A., R. FERGUS et W. T. FREEMAN (2008). « 80 million tiny images : a large dataset for non-parametric object and scene recognition ». Dans : *IEEE Transactions on Pattern Analysis and Machine Intelligence* 30.11.

VAPNIK, Vladimir N (1999). « An overview of statistical learning theory ». Dans : *Neural Networks, IEEE Transactions on* 10.5, p. 988–999.

VAPNIK, Vladimir N. (1995). *The nature of statistical learning theory*. New York, NY, USA : Springer-Verlag New York, Inc. ISBN : 0-387-94559-8.

Victoria and Albert Museum. <http://www.vam.ac.uk/content/articles/h/history-of-fashion-1840-1900/>. Dernier accès août 2015.

WAEGEMAN, Willem et Luc BOULLART (2006). « An ensemble of Weighted Support Vector Machines for Ordinal Regression ». Dans : *International Journal of Computer Systems Science and Engineering* 3.1.

WEAVER, Gawain et Zach LONG (2009). « Chromogenic characterisation : a study of Kodak color prints, 1942–2008 ». Dans : *PMG Winter Meeting*.

WEIKUM, G., N. NTARMOS, M. SPANIOL, P. TRIANTAFILLOU, A. BENCZÚR, S. KIRKPATRICK, P. RIGAUX et M. WILLIAMSON (2011). « Longitudinal Analytics on Web Archive Data : It’s About Time ! » Dans : *Proceedings of the 5th biennial Conference on Innovative Data Systems Research (CIDR), Asilomar, CA, USA, January 9-12*, p. 199–202.

WEYAND, T., C.-Y. TSAI et B. LEIBE (2015). « Fixing WTFs : Detecting Image Matches caused by Watermarks, Timestamps, and Frames in Internet Photos ». Dans : *IEEE Winter Conference on Applications of Computer Vision (WACV’15)*.

XIAO, Jianxiong, James HAYS, Krista A. EHINGER, Aude OLIVA et Antonio TORRALBA. « Sun database : Largescale scene recognition from abbey to zoo ». Dans : *In CVPR*.

XU, Jinxi et W. Bruce CROFT (1996). « Query Expansion Using Local and Global Document Analysis ». Dans : *Proceedings of the 19th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval. SIGIR '96. Zurich, Switzerland : ACM*, p. 4–11. ISBN : 0-89791-792-8. DOI : 10.1145/243199.243202. URL : <http://doi.acm.org/10.1145/243199.243202>.

YAO, Y. Y. (mar. 1995). « Measuring retrieval effectiveness based on user preference of documents ». Dans : *J. Am. Soc. Inf. Sci.* 46.2, p. 133–145. ISSN : 0002-8231. DOI : 10.1002/(SICI)1097-4571(199503)46:2<133::AID-ASI6>3.0.CO;2-Z. URL : [http://dx.doi.org/10.1002/\(SICI\)1097-4571\(199503\)46:2<133::AID-ASI6>3.0.CO;2-Z](http://dx.doi.org/10.1002/(SICI)1097-4571(199503)46:2<133::AID-ASI6>3.0.CO;2-Z).

YUWA, Hedrick-Wong et Choog DESMOND (2013). *MasterCard :Global Destination Cities Index*. Rap. tech. MasterCard Worldwide Insights, p. 41.

ZHANG, Haipeng, Mohammed KORAYEM, David J CRANDALL et Gretchen LEBUHN (2012). « Mining photo-sharing websites to study ecological phenomena ». Dans : *Proceedings of the 21st international conference on World Wide Web. ACM*, p. 749–758.

ZHANG, Yating, Adam JATOWT et Katsumi TANAKA. « Search for Images of Historical Objects Using Wikipedia ». Dans :

Annexe A

Frises chronologiques

A.1 Frise chronologique représentant l'évolution du temps en géologie

Nous pouvons observer dans cette frise¹ chronologique que les périodes temporelles s'amenuisent au fil du temps. Plus l'information est récente plus elle est précise et détaillée.

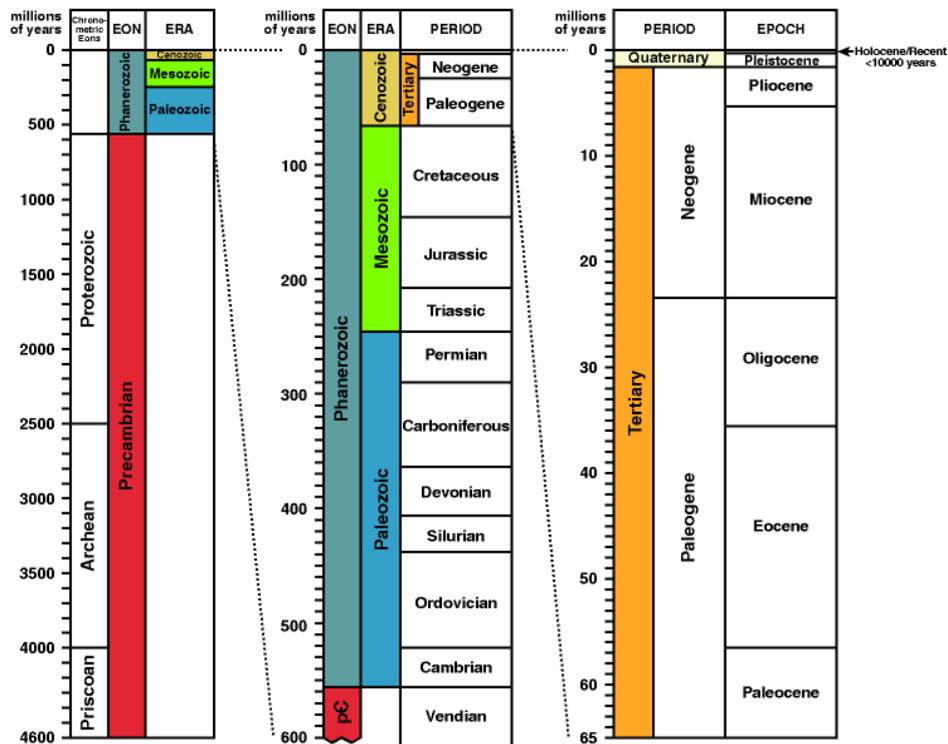


ILLUSTRATION A.1: Frise chronologique géologique. (source talkorigins.org)

1. <http://www.talkorigins.org/faq/timescale.html> dernier accès août 2015

A.2 Frises chronologiques retraçant l'évolution de la photographie



ILLUSTRATION A.2: Première photographie « Point de vue du Gras », depuis la fenêtre de la maison de Joseph Nicéphore Niépce. (support ayant imprimé la lumière). (source Harry Ransom Humanities Research Center, The University of Texas, Austin)

TABLEAU A.1 Frise chronologique retraçant l'évolution de la photographie [1826, 1888].

1826	• première photographie « Point de vue du Gras », depuis la fenêtre de la maison de Joseph Nicéphore Niépce (Camera obscura) c.f. illustration A.2
1837	• premier appareil permettant de réaliser des daguerréotypes
1839	• premier « selfie » pris par l'américain Robert Cornelius [1809, 1893] c.f. illustration A.3
1847	• première photographie de presse
1861	• première photographie couleur prise par Thomas Sutton c.f. illustration A.4

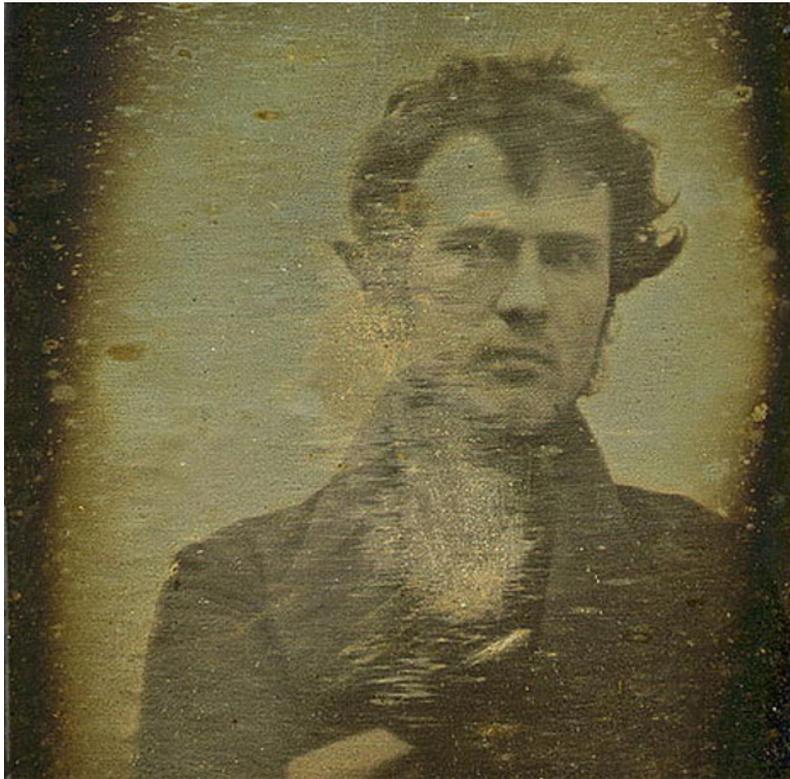


ILLUSTRATION A.3: Premier selfie pris par l'américain Robert Cornelius [1809, 1893] en 1839.



ILLUSTRATION A.4: Première photographie couleur prise par Thomas Sutton en 1861.

TABLEAU A.2 *Frise chronologique retraçant l'évolution de la photographie* [1888, 1957].

1888	•	premier appareil permettant de réaliser des photographies sur film souple (le <i>Kodak</i> 100 vues)
1889	•	avec le slogan « vous appuyez sur le bouton, nous faisons le reste », George Eastman a commercialisé le premier appareil photographique portable, en 1888. « You press the button we do the rest » (publicité <i>Kodak</i> , 1889)
1900	•	<i>Kodak Brownie</i> le premier appareil photographique commercialisé permettant de recharger les films soi-même
1915	•	stylo <i>Kodak Autographique</i> https://en.wikipedia.org/wiki/Autographic_film
1925	•	l'un des premiers appareils effectuant des photographies en 35mm, le <i>LEICA I</i>
années 1930	•	arrivée des appareils photographiques réflexes
1935	•	arrivée de la pellicule couleur <i>Kodachrome</i>
1948	•	premier appareil photographique instantané : <i>Polaroid</i> modèle 95



ILLUSTRATION A.5: Première photographie digitale (scannée), cette image représente le fils de Russell Kirsch, l'inventeur du scanner à plat en 1957.

TABLEAU A.3 Frise chronologique retraçant l'évolution de la photographie [1957, 2000].

1957	• première photographie digitale (scannée), cette image représente le fils de Russell Kirsch l'inventeur du scanner à plat <i>c.f.</i> illustration A.5
1963	• <i>Instamatic</i> de <i>Kodak</i> avec son slogan : « clic clac merci <i>Kodak</i> »
dans les années 1970	• les premiers scanners à plat apparaissent dans les années 1970
décembre 1975	• premier appareil photographique numérique construit par un ingénieur de <i>Kodak</i> (enregistrant sur cassette audio)
au cours des années 1980	• sortie des systèmes d'horodatage sur la face avant des photographies <i>e.g.</i> <i>canon quartz date back E</i> , <i>Nikon Quartz date</i> , <i>CHINON : INFO BACK-2 (DP-520)</i> , <i>INFO BACK (DB-010)</i> , <i>Auto Date (AD-510)</i> etc.
années 1980	• arrivée des appareils photographiques 35mm « Visez et déclenchez »
1991	• première version IPTC/IIM : <i>International Press and Telecommunications Council (IPTC)/ Information Interchange Model (IIM)</i>
1991	• premier réflexe numérique mono-objectif professionnel (<i>Kodak DCS 100</i>)
octobre 1995	• première version EXIF : <i>Exchangeable Image File Format</i>
fin des années 1990	• le boom de la photographie numérique à la fin des années 90
2000	• premier téléphone portable avec appareil photographique intégré (le <i>sharp j-sh04</i>)
avril 2001	• première version d'XMP : <i>eXtensible Metadata Platform</i>
29 juin 2007	• sortie du premier modèle d' <i>iPhone</i> par <i>Apple</i>
2007	• faillite de <i>Kodak</i>
2010 ?	• Applications pour ajouter des horodatages sur les images (<i>Google play</i>) https://play.google.com/store/apps/details?id=com.teapps.photostamperfree
fin des années 2000	• ajout des données GPS (arrivée des modules à la fin des années 2000 : autour de 2008)

Annexe B

Impacts du numérique sur l'économie de la photographie

B.1 Déclin des ventes de pellicules et d'appareils

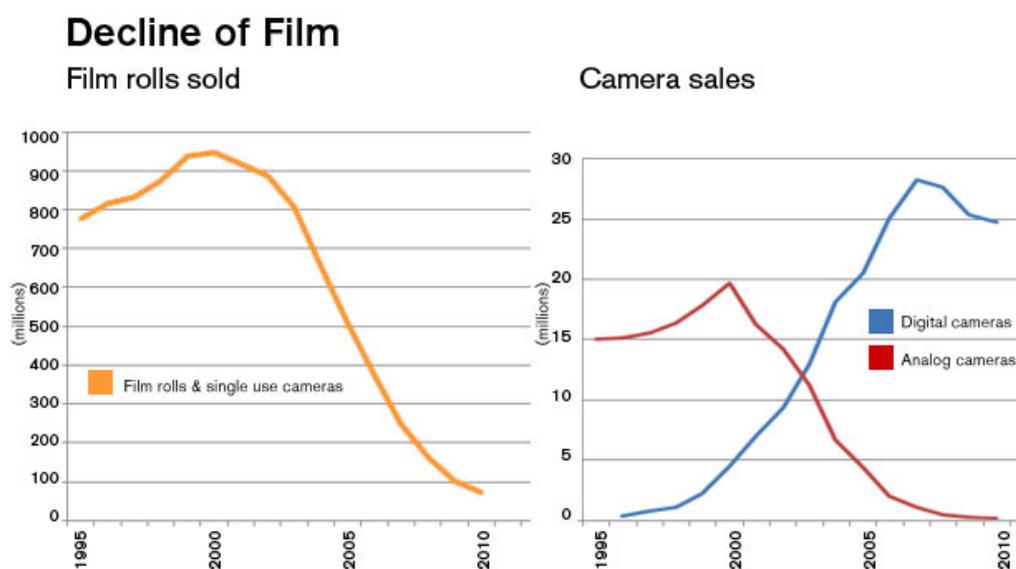


ILLUSTRATION B.1: Le déclin de la pellicule. (source technologyreview.com)

On peut voir dans le graphique présent dans l'illustration B.1 sur la gauche les « ventes de pellicules et d'appareils jetables en jaune » entre les années 1995 et 2010. Sur la droite les « ventes d'appareil photographique (argentique en rouge et numérique en bleu) » entre les années 1995 et 2010.

B.2 Déclin des développements de pellicules

On peut voir dans le graphique présent dans l'illustration B.2 sur la gauche les « Quantité d'impressions de photographies (des consommateurs Nord Américains) (développement des pellicules en rouge et impression des photographies numériques en bleu) » entre 2002 et 2011. Sur la droite les « Dépenses pour les impressions de photographies (des consommateurs Nord Américains) (pellicule en rouge et numérique en bleu) » entre 2004 et 2011.

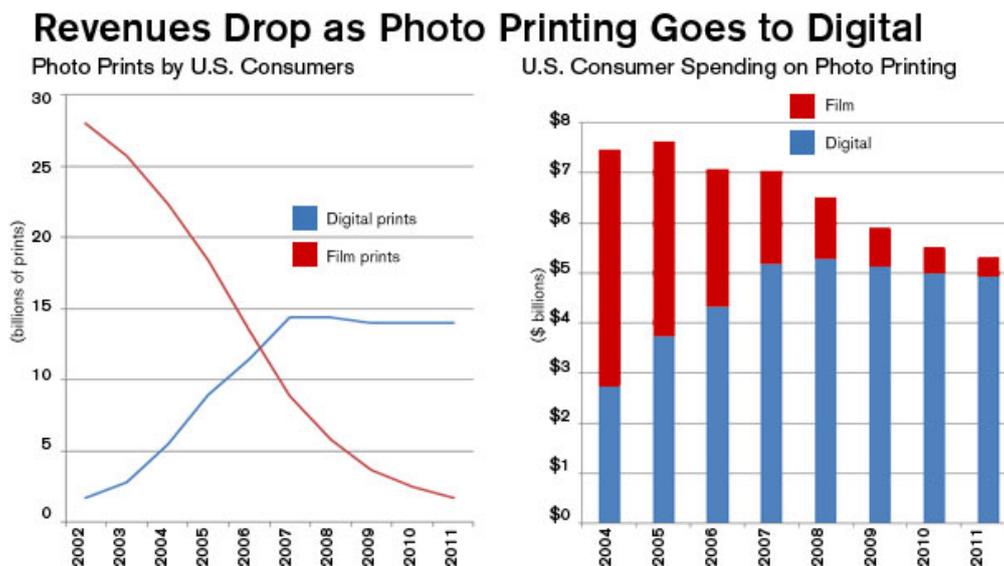


ILLUSTRATION B.2: La chute du chiffre d'affaires des impressions avec l'arrivée de la photographie en numérique. (source technologyreview.com)

B.3 Faillite de Kodak en 2012

Un graphique montrant l'évolution de l'action de *Kodak* en bourse. Le monde digital a été extrêmement néfaste pour *Kodak* qui était une des entreprises les plus prospères et qui plus est, était pionnière dans l'histoire de la photographie. Le comble de l'histoire étant que le premier appareil photographique digital ait été développé par un ingénieur de chez *Kodak* en décembre 1975.

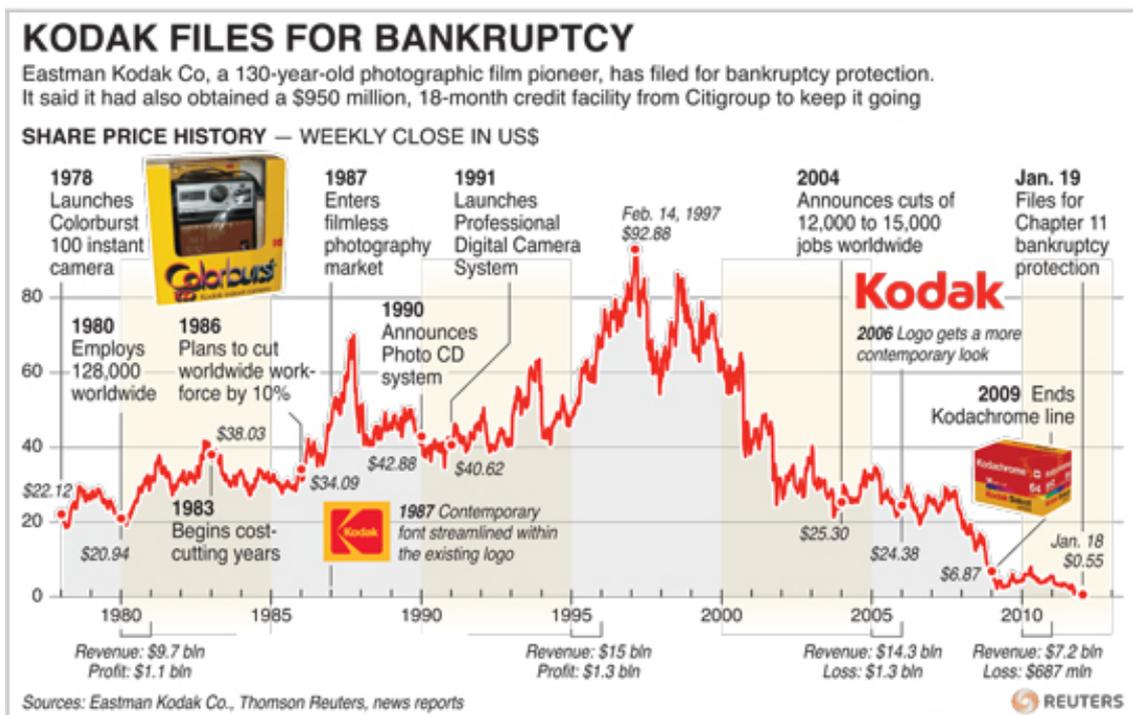


ILLUSTRATION B.3: Kodak dépose le bilan en 2012. (source reuters.com)

Annexe C

Spécifications de la norme XMP

Ceci est un extrait¹ détaillé des spécifications XMP² d'Adobe concernant les données temporelles.

XMP Schema Definitions

C.1 Category : Schema properties are internal or external

- Internal metadata must be maintained by an application. It can include system-level information (such as modification date) or information that an editing application has access to (such as the number of words in a document). An example is `xmp:ModifyDate`. Users should not be allowed to change the values of such properties. When a file is saved, an application should provide valid values for all internal properties. If an application does not set the value of an internal property, it should discard any value that may have existed previously.
- External metadata must be set by a user, and is independent of the contents of the document. External modifications should be displayed by the editing application but are not acted upon. Unless changed by the user, external properties are preserved on output. An example is `dc:creator`.

1. Ce document est disponible en version complète à l'adresse suivante : <https://partners.adobe.com/public/developer/en/xmp/sdk/XMPspecification.pdf>.

2. XMP : *Extensible Metadata Platform* – plate-forme de métadonnées évolutive

C.2 EXIF Dates

All date/time values are stored in XMP using ISO 8601 format. This is a combined date and time, with fractional seconds, and a time zone designation. The binary EXIF values generally separate the fractional seconds. EXIF 2.1 lacks time zone information ; this has been partially added in EXIF 2.2. When converting to XMP, the fractional seconds should be included. If no time zone is contained in the EXIF, convert to XMP assuming a local time.

C.3 Dublin Core Schema Date

dc :date (External) Date(s) that something interesting happened to the resource.

C.4 XMP Basic Schema

xmp :CreateDate (Internal) The date and time the resource was originally created

xmp :MetadataDate (Internal) The date and time that any metadata for this resource was last changed. It should be the same as or more recent than xmp :ModifyDate.

xmp :ModifyDate (Internal) The date and time the resource was last modified.
NOTE :The value of this property is not necessarily the same as the file's system modification date because it is set before the file is saved.

C.5 Photoshop Schema

photoshop :DateCreated (External) The date the intellectual content of the document was created (rather than the creation date of the physical representation), following IIM conventions. For example, a photo taken during the American Civil War would have a creation date during that epoch (1861-1865) rather than the date the photo was digitized for archiving.

C.6 XMP Schemas (EXIF Schemas)

C.6.1 tiff prefix schema

tiff :DateTime (Internal) TIFF tag 306, 0x132 (primary) and EXIF tag 37520, 0x9290 (subseconds). Date and time of image creation (no time zone in EXIF), stored in ISO 8601 format, not the original EXIF format. This property includes the value for the EXIF SubSecTime attribute. NOTE :This property is stored in XMP as xmp :ModifyDate.

C.6.2 exif prefix schema

exif :DateTimeOriginal (Internal) EXIF tags 36867, 0x9003 (primary) and 37521, 0x9291 (subseconds). Date and time when original image was generated, in ISO 8601 format. Includes the EXIF SubSecTimeOriginal data.

exif :DateTimeDigitized (Internal) EXIF tag 36868, 0x9004 (primary) and 37522, 0x9292 (subseconds). Date and time when image was stored as digital data, can be the same as DateTimeOriginal if originally stored in digital form. Stored in ISO 8601 format. Includes the EXIF SubSecTimeDigitized data.

exif :GPSTimeStamp (Internal) GPS tag 29 (date), 0x1D, and, and GPS tag 7 (time), 0x07. Time stamp of GPS data, in Coordinated Universal Time. NOTE :The GPSTimeStamp tag is new in EXIF 2.2. The GPS timestamp in EXIF 2.1 does not include a date. If not present, the date component for the XMP should be taken from exif :DateTimeOriginal, or if that is also lacking from exif :DateTimeDigitized. If no date is available, do not write exif :GPSTimeStamp to XMP.

Annexe D

Véracité des données EXIF et XMP extraites à partir de 150 000 images

Nous avons dans le chapitre 5 collecté un ensemble de 150 000 pages web traitant de villes obtenus par l'usage de requêtes temporelles. Nous reportons ci-dessous les statistiques correspondantes aux métadonnées numériques extractibles depuis les fichiers numériques. Pour obtenir les informations EXIF nous avons lu le contenu du champ `DateTimeOriginal` s'il n'était pas présent nous avons utilisé le champ `DateTime`. Dans le cas où ce contenu était présent nous avons extrait les 4 premiers caractères qui représentent en principe l'année. Pour obtenir les informations XMP nous avons exploité la propriété `photoshop:DateCreated` à l'espace de nom¹ : `http://ns.adobe.com/photoshop/1.0/` lorsqu'elle était déterminée. De la même manière que pour les données EXIF nous avons extrait les 4 premiers caractères qui, en théorie, doivent eux aussi représenter une année. Nous reportons dans l'annexe C plus d'informations sur les données XMP.

Le graphique de l'illustration D.1 représente la distribution des photographies entre [1899, 2016] à partir des données extraites des informations EXIF. Les données EXIF en dehors de la période [1899, 2016] sont reportées dans le tableau D.1. Le graphique dans l'illustration D.2 représente la distribution des photographies entre [1851, 2016] à partir des données extraites des informations XMP. Les données XMP en dehors de la période [1851, 2016] sont reportées dans le tableau D.2.

1. En anglais : *Namespace*

D.1 Statistiques données EXIF et XMP

Comme nous pouvons l'observer, dans le cadre des contenus anciens, les informations de temporalité en provenance des métadonnées numériques ne sont pas fidèles à la réalité. Le but que nous nous sommes fixés est donc d'outre-passer cette réalité qui rend les données traitant de sujets anciens inexploitable à leurs justes valeurs.

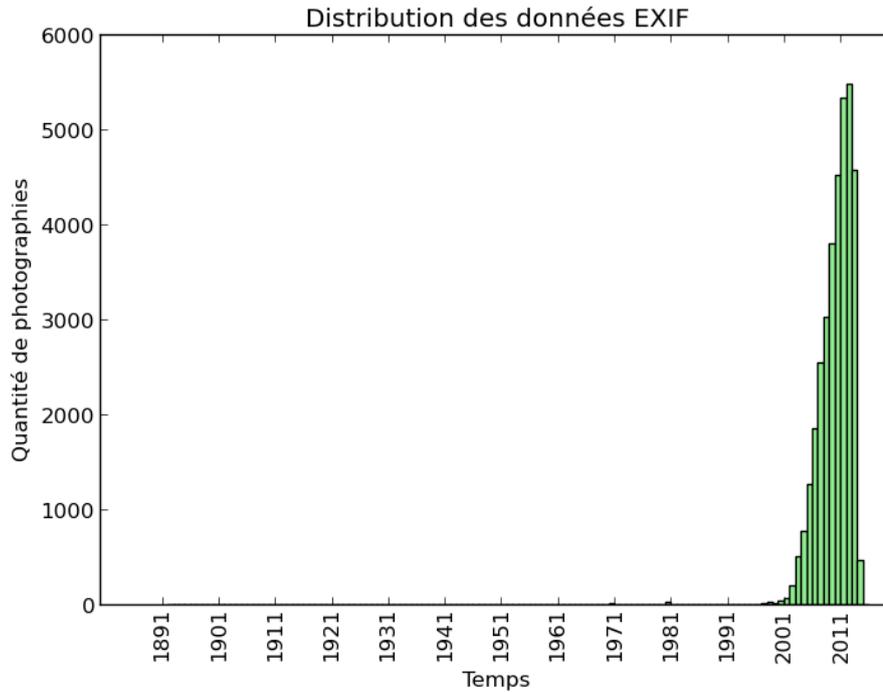


ILLUSTRATION D.1: Images datées dans l'intervalle [1899, 2016] selon les données EXIF extraites à partir des 150 000 images recueillies lors de la conception de notre jeu de données web, voir chapitre 5.

Informations EXIF	Nombre d'occurrences
0	234
20	1
611	1
1006	3
1204	1
1205	1
2020	1
2025	1
2028	4
2032	1
2033	1
2036	5
2158	6
2165	1
2212	1
2216	2
2243	1
4501	8
7011	1
7071	1
9082	1
Caractère non UNICODE	82
" " (1 espace)	8
" " (4 espaces)	122
0-00	1
04-0	1
05-0	1
07/0	1
1 :53	1
2 :20	1
May	1
Mon	2
None	114 971
Sun	1
Thu	2
Total Résultat	150 093

TABLEAU D.1: Données EXIF hors années [1899, 2016] extraites à partir des 150 000 images recueillies lors de la conception de notre jeu de données web, voir chapitre 5.

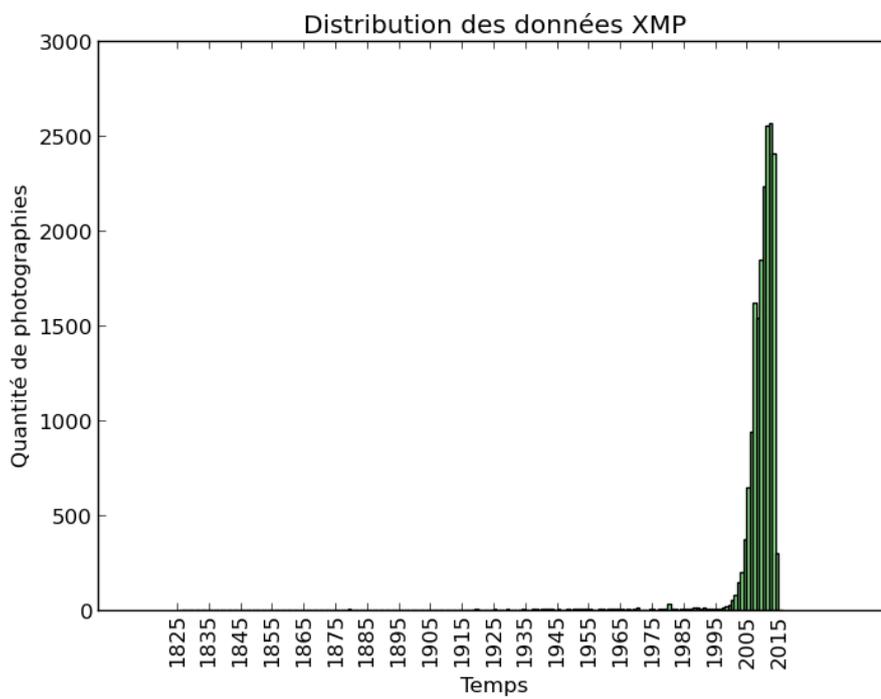


ILLUSTRATION D.2: Images datées dans l'intervalle [1851, 2016] selon les données XMP extraites à partir des 150 000 images recueillies lors de la conception de notre jeu de données web, voir chapitre 5.

Informations XMP	Nombre d'occurrences
0	302
612	1
701	1
703	1
704	8
705	3
706	1
708	1
712	1
803	1
808	1
811	2
1006	3
1007	1
1009	3
1011	1
1101	2
1753	1
2025	1
2028	4
2032	1
2033	1
2036	5
2063	1
2077	2
2158	6
2165	1
2212	1
2216	2
2243	1
4501	8
" " (4 espaces)	24
0-00	11
01 D	1
03 J	1
05 O	1
1 :53	1
11/9	1
12/1	1
12/8	1
13 D	1
13 F	1
14 N	1
15 J	1
15 N	2
17 J	1
19 J	1
2 :20	1
2/12	1
26 A	1
4/22	1
6-00	1
9/29	1
None	131 636
Total Résultat	150 093

TABLEAU D.2: Données XMP hors années [1851, 2016] extraites à partir des 150 000 images recueillies lors de la conception de notre jeu de données web, voir chapitre 5.

Annexe E

Autres jeux de données ordinaux

Il existe de nombreux jeux de données ordinaux qui ont été utilisés dans la littérature. En voici une liste des plus connus.

E.1 Jeux de données naturellement ordinaux

Autres jeux de données de Weka

Weka HALL, FRANK, HOLMES, PFAHRINGER, REUTEMANN et WITTEN (2009) est un logiciel de fouille de données.

Pasture production

Le jeu de données *Pasture production*¹ contient 23 attributs qui sont réels, entiers ou encore des valeurs énumérées. Ce jeu a été utilisé par COSTA, ALONSO et CARDOSO (2008).

Jeux de données ordinaux de Weka : jeux du monde réel (ERA, ESL, LEV et SWD)

Voici leurs définitions en anglais :

Weka Ordinal, real-world datasets (ERA, ESL, LEV and SWD)

1. <http://www.cs.waikato.ac.nz/~ml/weka/agridatasets.jar> dernier accès août 2015

These 4 datasets ARIE (*Weka Ordinal, real-world datasets (ERA, ESL, LEV and SWD)*) have been donated by Dr. Arie Ben David to Weka HALL, FRANK, HOLMES, PFAHRINGER, REUTEMANN et WITTEN (2009) users, they are named ERA, ESL, LEV and SWD respectively for Employee Rejection

Acceptance, Employee Selection, Lecturers Evaluation and Social Workers Decisions. In these datasets all inputs and outputs data are numerical values and they are ordinal (also called monotone).

The ERA dataset is a collection of 1 000 instances of candidates applying to jobs. The input part of those instances describe with 4 numerical values the 4 most important qualities representing a candidate (past experience, verbal skills. . .). The output part is a subjective judgement expressed as a degree representing the candidate acceptance rate to the application. Degrees of acceptance are gradually ordered values from 1 to 9 (lowest means reject). This data set was collected during a MBA academic course.

The ESL dataset is a smaller collection of candidate applying to job offers (only 488 profiles). The input part is also composed of 4 attributes but those attributes are determined this time by experts after interviews with candidates, psychometric tests. . .The output part has the same range and meaning that the one of ERA but it's here an overall score.

The LEV dataset is composed of 1 000 anonymous lecturer evaluations. The features describing evaluations are 4 attributes such as : oral skills of the lecturer, contribution to professional/general knowledge of the students. . .The output is the total evaluation of the lecturer's performance given from 0 to 4.

The SWD dataset contains 1 000 assessments of qualified social workers. Those assessments are expressed by 10 attributes concerning the risk children have if they stay inside their families. The output value is the final evaluation risk for the children, this value ranges from 2 to 5.

Some or all of these 4 datasets have been used CARDOSO et SOUSA (2011) ; DUIVESTELJN et FEELDERS (2008) ; COSTA, ALONSO et CARDOSO (2008) ; COSTA, SOUSA et CARDOSO (2010) with variates ordinal methods.

Dépôt de l'UCI

Le site de l'UCI est une ressource importante pour les jeux de données d'apprentissage automatique et donc les jeux de données ordinaux. On trouve notamment parmi le *Machine Learning repository* les jeux suivants :

Three Aspects of Cognitive Development Balance Scale dataset

Le jeu de données *Balance Scale Data Set*² a été offert par Tim Hume ce jeu a été généré pour modéliser des expériences psychologiques qui ont été menées sur les aspects cognitifs par SIEGLER.

Il a été utilisé dans les travaux de CARDOSO et SOUSA (2011); KRAMER, WIDMER, PFAHRINGER et GROEVE (2001); COSTA, SOUSA et CARDOSO (2010) pour différentes tâches études de nouvelles méthodes basées sur les SVMs, les arbres de régression ou l'analyse de mesures d'évaluation.

Nursery³ et **Car Evaluation Data Set**⁴ ont été utilisés par KRAMER, WIDMER, PFAHRINGER et GROEVE (2001).

Pima Indians Diabetes⁵, **Haberman Survival**⁶ et **Australian credit approval**⁷ Proviennent du dépôt de l'*UCI ML* ces jeux de données ont été utilisés par DUIVESTELJN et FEELDERS (2008).

Jeux de données de sources diverses

BCCT

Jeu de données BCCT introduit dans CARDOSO et CARDOSO (2007) utilisé par CARDOSO et SOUSA (2011); COSTA, SOUSA et CARDOSO (2010).

Windsor Housing

*Windsor Housing*⁸ introduit dans ANGLIN et GENÇAY (1996) c'est un jeu de données monotones représentant les caractéristiques de maisons ainsi que leurs prix dans la ville de Windsor dans l'Ontario au Canada. Ce jeu de données a été utilisé par DUIVESTELJN et FEELDERS (2008).

2. <http://archive.ics.uci.edu/ml/datasets/Balance+Scale> dernier accès août 2015
3. <http://archive.ics.uci.edu/ml/datasets/Nursery> dernier accès août 2015
4. <http://archive.ics.uci.edu/ml/datasets/Car+Evaluation> dernier accès août 2015
5. <http://archive.ics.uci.edu/ml/datasets/Pima+Indians+Diabetes> dernier accès août 2015
6. <http://archive.ics.uci.edu/ml/datasets/Haberman's+Survival> dernier accès août 2015
7. [http://archive.ics.uci.edu/ml/datasets/Statlog+\(Australian+Credit+Approval\)](http://archive.ics.uci.edu/ml/datasets/Statlog+(Australian+Credit+Approval)) dernier accès août 2015
8. <http://econ.queensu.ca/jae/1996-v11.6/anglin-gencay/> dernier accès août 2015

Biodegradability dataset

À l'origine, cette base de données provient du « *Handbook of environmental degradation rates* » « Handbook of environmental degradation rates » (1991) ensuite, une version alternative à été réalisé par DZEROSKI, BLOCKEEL, KOMPARE, KRAMER, PFAHRINGER et LAER (1999). Enfin KRAMER, WIDMER, PFAHRINGER et GROEVE (2001) ont discrétisé les valeurs présentes pour faire un jeu de données ordinal.

Mesh

Présenté et introduit dans DOLSAK et MUGGLETON (1992). KRAMER, WIDMER, PFAHRINGER et GROEVE (2001) ont utilisé ce jeu de données du fait de sa nature ordinaire.

OHSUMED

OHSUMED⁹ est un jeu composé de presque 400000 références d'articles écrits dans le domaine de la médecine. Ce jeu à été utilisé dans des tâches de recherche d'information et aussi en classification ordinaire dans les travaux de HARRINGTON (2003) ; HERBRICH, GRAEPEL, OBERMAYER et al. (1999).

Cystic Fibrosis Database

Cystic Fibrosis Database¹⁰ est un jeu de données contenant des documents publiés entre les années 1974 et 1979 parlant des aspects de la mucoviscidose, ce jeu a été utilisé par HARRINGTON (2003).

Gene expression analysis

Introduit par SINGH, FEBBO, ROSS, JACKSON, MANOLA, LADD, TAMAYO, RENSHAW, D'AMICO, RICHIE et al. (2002) ce jeu de données a été utilisé en classification ordinaire par CHU et GHAHRAMANI (2004).

9. <http://davis.wpi.edu/xmdv/datasets/ohsumed.html> dernier accès août 2015

10. <http://people.ischool.berkeley.edu/~hearst/irbook/cfc.html> dernier accès août 2015

Annexe F

Captures d'écrans et présentations des bases d'images découvertes

Nous représentons dans cette section une série de captures d'écrans des différents sites sur lesquels nous avons découvert des bases d'images datées.

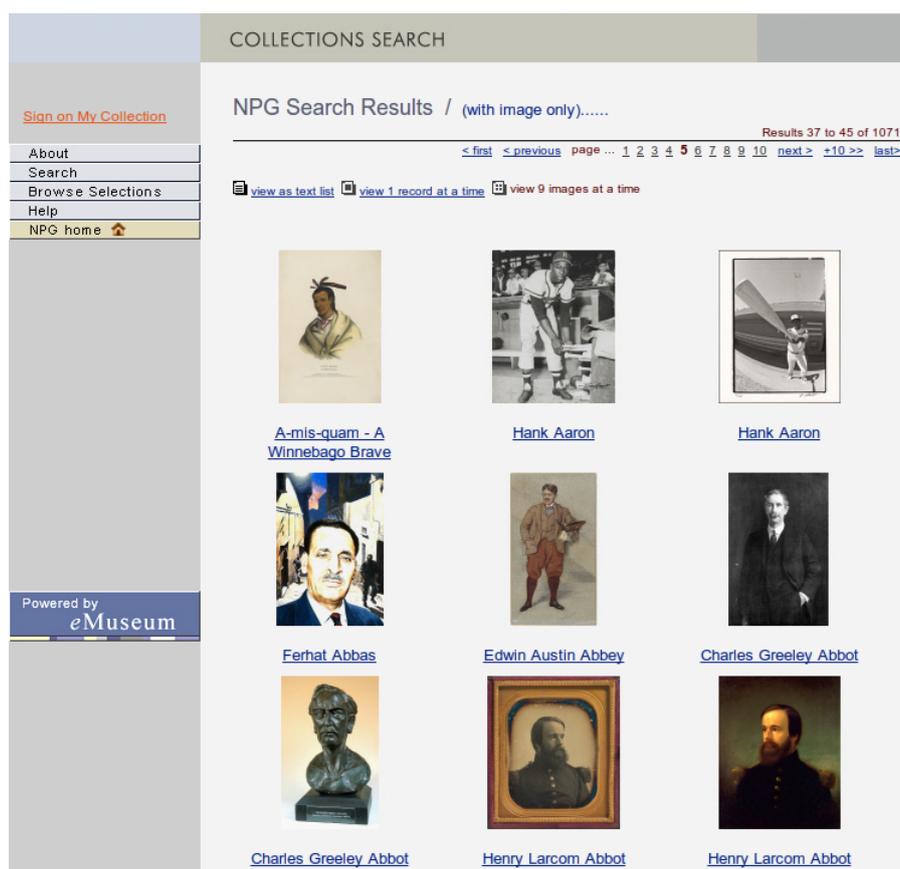


ILLUSTRATION F.1: Interface web de « National Portrait Gallery ».

 **Exposition des canots autos, Monaco, avril 1906**

Num. Phot. 3172 – image/jpeg



[Voir le document](#)



Auteur : [Trutat, Eugène \(1840-1910\)](#), Photographe
date de publication: 1906

ILLUSTRATION F.2: Une photographie de la base de la bibliothèque de Toulouse.

Auteur : [Trutat, Eugène \(1840-1910\)](#), Photographe
date de publication: 1906
Desc. Mat. : 9 x 12 cm
Type : [image fixe](#)
[photographie](#)
[Photographies](#)
[Négatifs sur plaque de verre](#)
[Noir et blanc](#)
[Scènes de rue](#)
[Goerz-Anschutz](#)
[Jougla-Lumière](#)
Note : Note manuscrite de Trutat : " Monaco : exposition des canots auto, avril 1906, Goerz-Anschutz, Jougla b. bleue ".
Cote : Bibliothèque municipale de Toulouse, TRU C 2021
URL : http://numerique.bibliotheque.toulouse.fr/ark:/74899/B315556101_TRUC2021
Sujet : [Rue](#)
[Automobile](#)
[Fiacre](#)
[Cheval](#)
[Rail](#)
[Pavement](#)
[Drapeau](#)
[Fête](#)
Lieu : [Monaco \(Principauté de\)](#)
Doc. associé : [TRU C 1846 Exposition des canots autos, Monaco, avril 1905](#)
[TRU C 2032 Exposition des canots autos, Monaco, avril 1905](#)
Provenance : Toulousains de Toulouse et Amis du Vieux Toulouse
Document http://catalogues.toulouse.fr/web2/tramp2.exe/do_keyword_search

ILLUSTRATION F.3: Détails disponibles pour une photographie de la base de la bibliothèque de Toulouse.

The screenshot shows the Gallica website interface. At the top, there is a search bar with the text 'Tout Gallica' and a search button labeled 'OK >> Recherche avancée'. Below the search bar, the page indicates '224 082 résultats page 7 sur 14 939'. On the left side, there is a sidebar with filters: 'Type de document' (Image: 224 082), 'Auteur' (1 000 premiers résultats), 'Date d'édition' (XXe siècle: 98 674, XIXe siècle: 51 997, XVIIIe siècle: 25 621, XVIIe siècle: 7 489, XVIe siècle: 3 192), and 'Langue' (Français: 99 591, fr: 10 574, zxx: 2 763, Anglais: 1 293, Latin: 1 171). The main content area displays a list of search results. Each result includes a thumbnail image, a title, and a brief description. The first result is '[Recueil. Albums de photographies de la famille Dupré]. 73 fotogr. [à l'origine dans un album ; les fotogr. ont été remontées sur des pl. 28 x 20 cm] -- Informations détaillées'. The second result is '[Recueil. Albums de photographies de la famille Dupré]. 19 fotogr. - 16 x 23 cm -- Informations détaillées'. The third result is '[Recueil.] Château de Gaesbeck près Bruxelles (Brabant méridional) : à la Marquise Arconati Visconti -- 1900-1920 Informations détaillées'. The fourth result is 'Figaro-Album. Photographies de différents personnages français appartenant à la politique, aux arts, aux sciences, au théâtre, etc., etc -[s.n.]-1875 Informations détaillées'. Each result also includes a link to 'Image(s)' and options to 'Voir tous les volumes du même ensemble éditorial' and 'Accéder au document'.

ILLUSTRATION F.4: Présentation base Gallica, de nombreux supports différents recueils et albums

BnF Bibliothèque nationale de France | accessibilité | English | Español | Português | Deutsch | aide | Une question ? | Espace personnel

gallica BIBLIOTHÈQUE NUMÉRIQUE

Tout Gallica Livres Manuscrits Cartes Images Presse et revues Paroles et musiques Partitions

OK >> Recherche avancée

> Accueil > 224 082 résultats page 7 sur 14 939

224 082 résultats Page 7 sur 14 939 15 résultats par page

Affichage : Trier par : Pertinence

Affiner

Ma recherche initiale Recherche avancée:

Limité à :

Type de document : - Image

Rechercher dans ces résultats

Type de document

Image (224 082)

Auteur (1 000 premiers résultats)

Lhéritier (1809-1885). Dessinateur (60)

Atget, Eugène (1857-1927). Photographe (20)

Dassier, Jean (1676-1763). Médailleur (16)

Sève, Jacques de. Dessinateur (10)

Vallotton, Félix (1865-1925). Graveur (9)

Date d'édition

XXe siècle (98 674)

XIXe siècle (51 997)

XVIIIe siècle (25 621)

XVIIe siècle (7 489)

XVIe siècle (3 192)

Langue

Français (99 591)

fr (10 574)

zxx (2 763)

Anglais (1 293)

Latin (1 171)

91 **[Recueil. Albums de photographies de la famille Dupré]. 73 fotogr. [à l'origine dans un album ; les fotogr. ont été remontées sur des pl. 28 x 20 cm]** -- Informations détaillées



Notice complète

Titre : [Recueil. Albums de photographies de la famille Dupré]. 73 fotogr. [à l'origine dans un album ; les fotogr. ont été remontées sur des pl. 28 x 20 cm]

Type : image fixe

Format : 2 vol. fotogr. en sépia

Format : image/jpeg

Droits : domaine public

Identifiant : ark:/12148/btv1b8432358r

Source : Bibliothèque nationale de France, département Estampes et photographie, 4-NA-417(A)

Relation : Notice d'ensemble : <http://catalogue.bnf.fr/ark:/12148/cb-40355548m>

Relation : <http://catalogue.bnf.fr/ark:/12148/cb-40355548m>

Provenance : bnf.fr

Date de mise en ligne : 18/01/2011

Fermer la notice complète

Image(s)

Voir tous les volumes du même ensemble éditorial

Accéder au document Feuilleter

92 **[Recueil. Albums de photographies de la famille Dupré]. 19 fotogr. - 16 x 23 cm** -- Informations détaillées



Image(s)

Voir tous les volumes du même ensemble éditorial

Accéder au document Feuilleter

93 **[Recueil.1 Château de Gaesbeck près Bruxelles (Brabant méridional) : à la**

ILLUSTRATION F.5: Présentation base Gallica (suite) la notice complète d'un album de photographies

La Médiathèque de l'Architecture et du Patrimoine
 IMAGES Base Mémoire : Archives photographiques
 50 / 512581 [Affiner la recherche](#) [Autre recherche](#) [Commande photo](#)

	☰	<p>Localisation France ; Auvergne ; Haute-Loire ; Léotoing</p> <p>Nom de l'édifice Château de Léotoing</p> <p>Nom de l'objet</p> <p>Légende <i>vue générale, ensemble Sud-Est</i></p> <p>Photographe Genestoux, Franck</p> <p>Date prise vue 2009</p> <p>Crédit photo Ministère de la Culture (France), Médiathèque de l'architecture et du patrimoine, Diffusion RMN</p>
	☰	<p>Localisation France ; Auvergne ; Haute-Loire ; Léotoing</p> <p>Nom de l'édifice Maison de village, ancien presbytère</p> <p>Nom de l'objet</p> <p>Légende <i>vue partielle, façade Est</i></p> <p>Photographe Genestoux, Franck</p> <p>Date prise vue 2009</p> <p>Crédit photo Ministère de la Culture (France), Médiathèque de l'architecture et du patrimoine, Diffusion RMN</p>
	☰	<p>Localisation France ; Auvergne ; Haute-Loire ; Léotoing</p> <p>Nom de l'édifice Eglise Saint-Vincent</p> <p>Nom de l'objet</p> <p>Légende <i>détail des corbeaux de pierre sculptés, façade Sud</i></p> <p>Photographe Genestoux, Franck</p> <p>Date prise vue 2009</p> <p>Crédit photo Ministère de la Culture (France), Médiathèque de l'architecture et du patrimoine, Diffusion RMN</p>
	☰	<p>Localisation France ; Auvergne ; Haute-Loire ; Léotoing</p> <p>Nom de l'édifice Château de Léotoing</p> <p>Nom de l'objet</p> <p>Légende <i>vue générale, ensemble Nord-Est</i></p> <p>Photographe Genestoux, Franck</p> <p>Date prise vue 2009</p> <p>Crédit photo Ministère de la Culture (France), Médiathèque de l'architecture et du patrimoine, Diffusion RMN</p>

ILLUSTRATION F.6: Présentation de la base de la Médiathèque.

Rechercher

Photos | Groupes | Personnes

The Commons

RECHERCHER

- Afficher le contenu provenant des organismes publics
- Afficher uniquement les photos
- Prise entre le 01/01/1826 et le 12/31/1998

Tri : Pertinent | Récent | Intéressant

Vue : Petit | Moyen | Détail | Diaporama



De The...



De The...



De National...



De The...



De The...



De The...



De State...



De Swedish...



De State...



De The...



De The...



De The...



De The...



De George...



De Swedish...



De The...

ILLUSTRATION F.7: Images de la base flickr « The Commons ».

C'était Caen ?. Le 31 décembre 1925, Caen réveillonne sous les eaux

Caen - 03 Janvier écouter



Edit. P. Sirelle. rue Hannon - Caen

Facebook Twitter 8 Google+ Achetez votre Journal numérique

Chaque samedi, cette rubrique met en relief une date clé de la semaine en cours. Le 31 décembre 1925, la crue de l'Orne engendre des inondations exceptionnelles.

« En deux heures, les quartiers de Vaucelles, de Saint-Jean et de Saint-Pierre étaient submergés littéralement. » *L'Ouest-Éclair*, ancêtre de *Ouest-France*, décrit la crue de l'Orne dans son édition du 1^{er} janvier 1926 comme « un véritable désastre ».



- 20/09 Foire de Caen. 60 000 visiteurs en trois jours, dont 30 000 ce dimanche
- 20/09 Foire de Caen. C'est reparti pour un dimanche sur les chapeaux de roue !
- 20/09 Journées du patrimoine à Caen. Classique, nouveau et moins connu
- 20/09 Ligue 1. Roland Courbis : "Je vais féliciter Patrice Garande"

Services Ouest-France

- Abonnés, découvrez vos privilèges
- Abonnés, gérez votre abonnement
- Achetez le journal numérique du jour (0,95€)
- Abonnez-vous à partir de 1€/mois
- La Boutique Ouest-France
- Jeux
- Annonces
- Obsèques / Dans nos cœurs

Les plus lus **Les plus commentés**

- Faits divers à Rennes. Des riverains sauvent une ado d'une agression - Faits divers
- Brest. Il étrangle sa mère... parce qu'elle passe l'aspirateur - Brest
- Bretagne. Deux veuves se disputent la résidence secondaire - Bretagne
- Miss France. La Miss Côtes-d'Armor est aussi... médecin ! - Bretagne
- Claire Chazal. Larmes et audience record pour

ILLUSTRATION F.8: Capture d'écran du site du journal Ouest-France représentant une rue de Caen photographiée le 31 décembre 1925 ainsi que son texte descriptif associé.

Annexe G

Informations complémentaires de notre jeu de données issu de *flickr*

G.1 Présentation d'un document de notre base

Nous présentons la structure des fichiers XML stockés à partir des données téléchargées depuis *flickr*. Voir l'illustration G.1 présentant un exemple d'image téléchargée depuis la base *flickr* « *The Commons* » et le contenu qui lui est associé que nous avons stocké dans un fichier XML, *c.f.* illustration G.2.

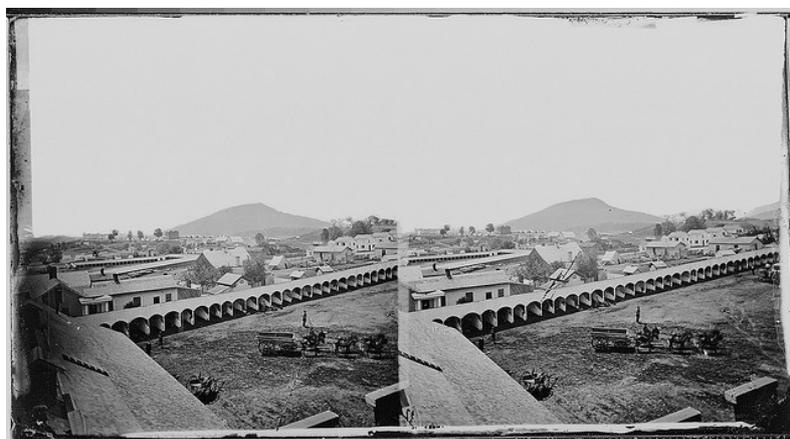


ILLUSTRATION G.1: Image 4152965237.jpg associée au fichier XML 4152965237.xml.
(source flickr « *The Commons* »)

G.1. Présentation d'un document de notre base

```
<photo id="4152965237"
      owner="35740357@N03"
      title="Fair grounds at Chattanooga, Tenn"
      description="
Original Caption: Fair grounds at Chattanooga, Tenn

U.S. National Archives' Local Identifier: 111-B-584

From:: Series: Mathew Brady Photographs of Civil War-Era Personalities
and Scenes, (Record Group 111)

Photographer: Brady, Mathew, 1823 (ca.) - 1896
Coverage Dates: ca. 1860 - ca. 1865

Subjects: American Civil War, 1861-1865
Brady National Photographic Art Gallery (Washington, D.C.)

Persistent URL: arcweb.archives.gov/arc/action/ExternalIdSearch?id=524992

Repository: Still Picture Records Section, Special Media Archives
Services Division (NWCS-S), National Archives at College Park, 8601 Adelphi
Road, College Park, MD, 20740-6001.

For information about ordering reproductions of photographs held by the
Still Picture Unit,
visit: www.archives.gov/research/order/still-pictures.html

Reproductions may be ordered via an independent vendor. NARA maintains a
list of vendors at
www.archives.gov/research/order/vendors-photos-maps-dc.html

Access Restrictions: Unrestricted
Use Restrictions: Unrestricted"
      dateupload="1259779784"
      datetaken="1863-01-01 00:00:00"
      datetakengranularity="8"
      ownername="The U.S. National Archives"
      originalformat="jpg"
      latitude="35.040713"
      longitude="-85.304203"
      accuracy="13"
      tags="civilwar mathewbrady usnationalarchives nara:arcid=524992"
      o_width="3000"
      o_height="1655"
      pathalias="usnationalarchives"/>
```

ILLUSTRATION G.2: Contenu du fichier XML 4152965237.xml associé à l'image 4152965237.jpg. (source flickr « The Commons »)

G.2 Présentation de notre système d'annotation d'images

Nous avons développé un système d'annotation qui nous a permis de sélectionner les photographies de rues parmi les images de notre base provenant de flickr. En rouge les images qui ne sont pas de rue pour l'annotateur réalisant le tri, en vert les photographies de rues et en blanc les images restantes à annoter par l'utilisateur.

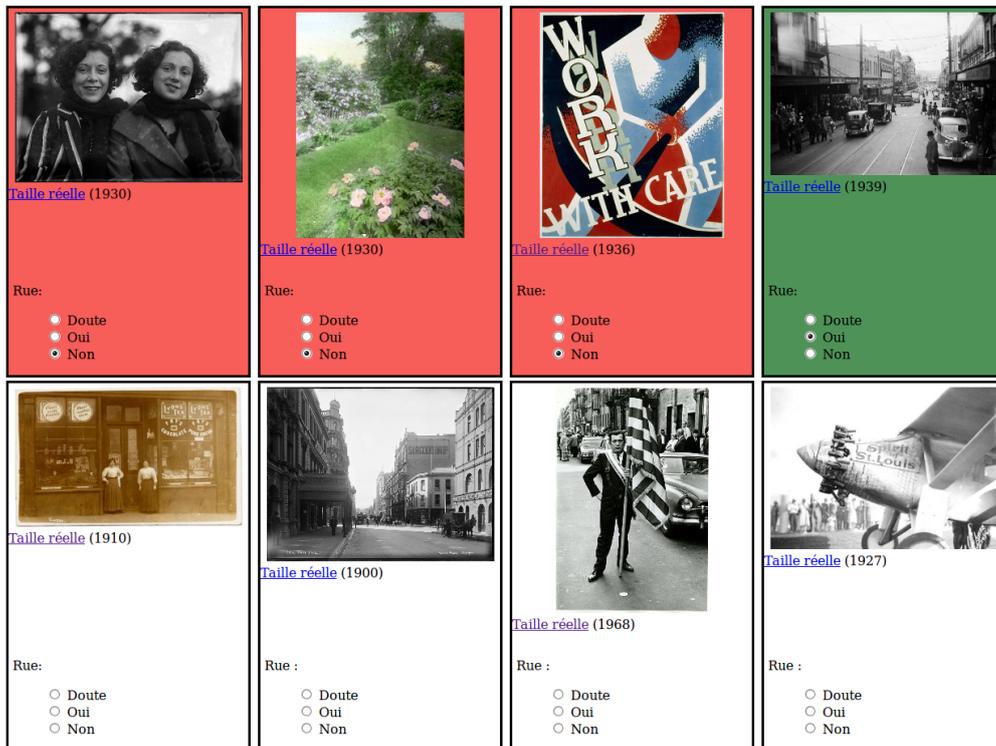


ILLUSTRATION G.3: Présentation de l'interface d'annotation réalisée.

G.3 Statistique de répartition des photographies

Comme nous le montrent les graphiques suivants G.4, G.5 et G.6 la répartition des images par année est fortement biaisée.

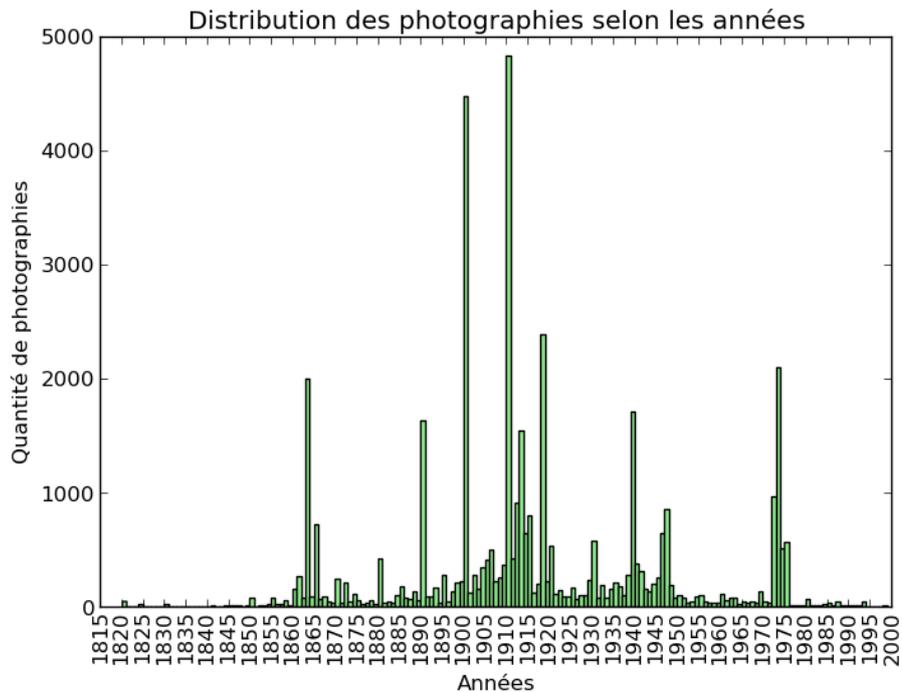


ILLUSTRATION G.4: Répartition par année de la collection complète des 42 310 documents numériques que nous avons extraits.

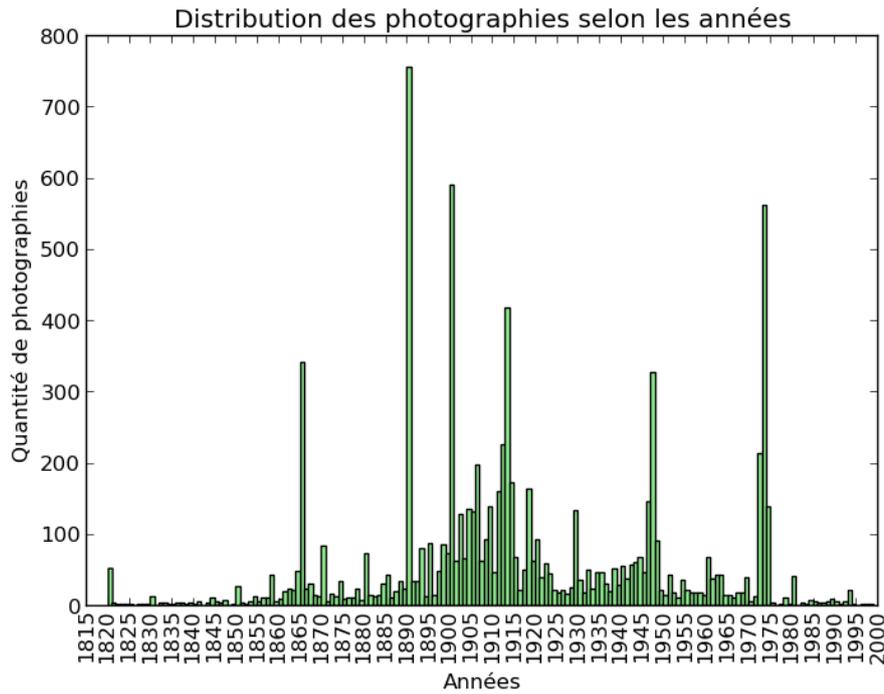


ILLUSTRATION G.5: Répartition par année de la collection des 8 831 documents numériques que nous avons extraits pour lesquels la précision est à l'année près.

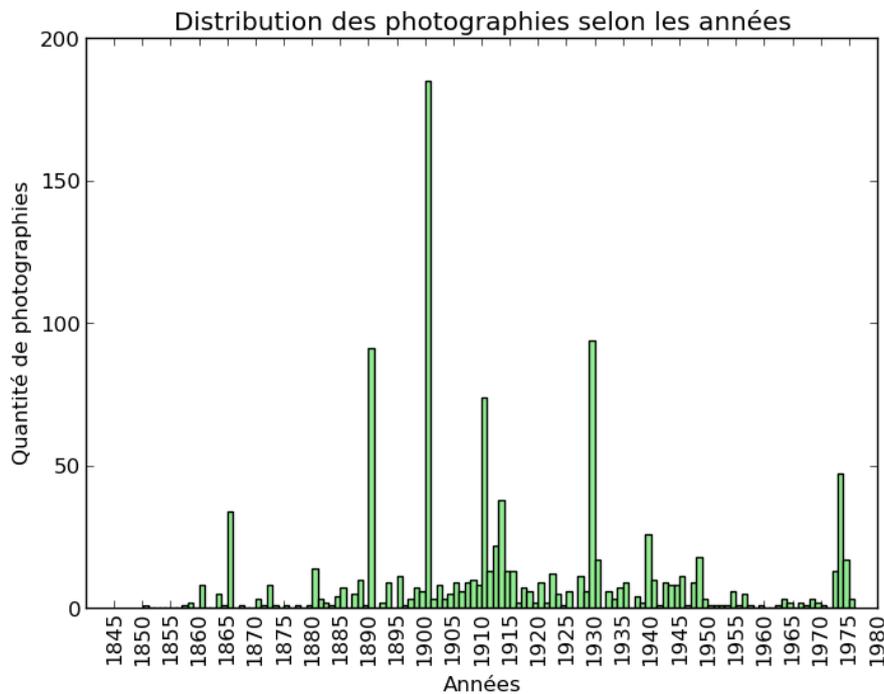


ILLUSTRATION G.6: Répartition par année de la collection 1 093 documents numériques que nous avons extraits et que nous avons annotés comme représentatifs de photographies de rues.

Annexe H

Informations complémentaires de notre jeu de données issu du web

H.1 Liste des 50 villes pour lesquelles nous avons effectué des recherches

Nous avons utilisé la liste des 50 villes les plus visitées durant l'année 2013. Ces données proviennent du *MasterCard Global Destination Cities Index 2013* (YUWA et DESMOND, 2013) et concernent 5 régions du globe. Pour chaque région, a été élu le « *Top 10* des villes dans lesquelles les visiteurs ont passé au moins une nuit¹ ». Nous reportons la liste de ces 5 régions et 50 noms de villes dans le tableau H.1. Nous avons utilisé les noms anglais de ces villes dans nos requêtes temporelles c'est pour cela que nous les reportons ainsi.

1. En anglais : *Top 10 Destination Cities by International Overnight Visitors*

H.1. Liste des 50 villes pour lesquelles nous avons effectué des recherches

Régions du globe	Villes
Asie et Pacifique	Bangkok
	Singapore
	Kuala Lumpur
	Hong Kong
	Seoul
	Shanghai
	Tokyo
	Taipei
	Beijing
	Guangzhou
Europe	London
	Paris
	Istanbul
	Barcelona
	Milan
	Rome
	Amsterdam
	Vienna
	Madrid
Prague	
Amérique Latine	Mexico City
	Buenos Aires
	Sao Paulo
	Lima
	San Jose
	Rio de Janeiro
	Bogotá
	Montevideo
	Quito
Caracas	
Moyen-Orient et Afrique	Dubai
	Riyadh
	Johannesburg
	Amman
	Lagos
	Cairo
	Abu Dhabi
	Tunis
	Casablanca
	Tel Aviv
Amérique du Nord	New York
	Los Angeles
	Miami
	Toronto
	Vancouver
	San Francisco
	Washington
	Chicago
	Montreal
Boston	

TABLEAU H.1: Liste des 50 villes au monde les plus visitées au monde pendant l'année 2013.
(source MasterCard)

H.2 Présentation d'un document de notre base

Nous avons utilisé des requêtes temporelles afin obtenir des documents source pour notre collection. Nous présentons les résultats obtenus avec la requête : Tokyo 1993 dans l'illustration H.1.

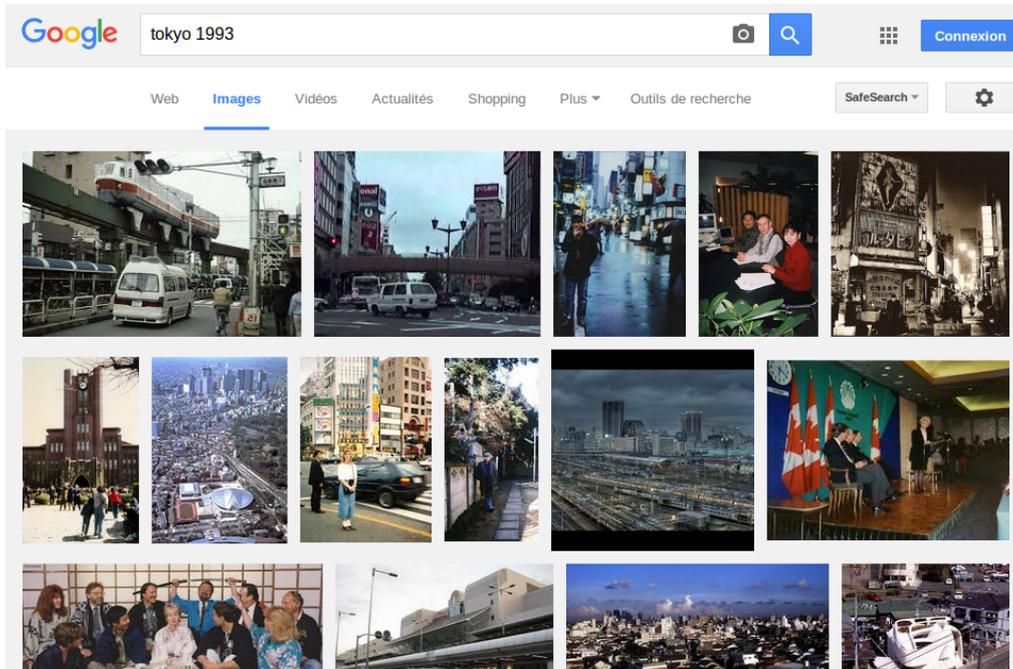


ILLUSTRATION H.1: Résultats de recherche avec le moteur de recherche Google images pour la requête Tokyo 1993.

Nous présentons un article de blog (c.f. illustrations H.2 et H.3) qui a été retrouvé grâce à cette requête. Cet article est un extrait de la page complète de *blog* qui est composée d'autres articles. Il a été écrit le 14 mars 2011, suite aux événements du 11 mars 2011 qui ont eu lieu au Japon. Nous avons récupéré les informations du *snippet* retourné par *Google images* (c.f. illustration H.4)² ainsi que l'image c.f. illustration H.5 et la page web complète afin que nous puissions l'annoter ultérieurement.

2. Notons qu'à l'époque de la réalisation de notre jeu de données, le document que nous présentons était classé à la première position, il est actuellement à la huitième position.

MONDAY, MARCH 14, 2011

Thoughts of Japan

It would be almost impossible to not recognize or be aware of the situation in Japan, at present. Even upon the writing of this, the story is still unfolding; it's a tragic story. One whose ramifications we will all feel in one way or another. However, my goal is not to focus on the doom and gloom that is and may lay ahead, but rather to remember a country that is one of my favorites to explore.



Postcard image of Mount Asama

In the summer of 1993, I spent 3 weeks in Japan visiting family friends. During this time we travelled to different regions of the country and spent many hours talking to Japanese college students who were eager to converse (in English) to Americans. It was a wonderful time for me, a Geography major, who loved cultural experiences. The people were gracious and hospitable. The older generation was more steeped in tradition, but the younger ones loved to talk politics and pop-culture. Former President Clinton was a hot topic at the time and the Japanese loved to ask questions about him.

ILLUSTRATION H.2: Début de l'article du blog. (source <http://reggies-words.blogspot.com/>)



Omotesando Dori, Tokyo, 1993.
Presently, 100,000 cars drive down this street daily.

Spending an extended period of time in another culture leaves its mark on you. I remember having thoughts of what Japan would be like, but nothing really prepared me for the modern, technologically advanced country that I experienced during my stay. The Land of the Rising Sun surprised me with its evergreen beauty, moonlit rice fields, majestic mountains, and miles of coastline (it is an island nation with over 6800 islands to its name) along with the uber-dense metropolis that is Tokyo and its surrounding prefectures. I was introduced to

new and interesting flavors in food (not your typical sushi) and was thoroughly impressed by its state of the art technology. In a funny sidenote, only in Japan have I seen and used a gold plated bathroom, including toilet that was outfitted with more features than I could think possible!

This country holds a fond spot in my heart. And right now, especially, I'm praying for its people.

-Reggie

Posted by Reggie's Words at 11:01 PM 3 comments:

  Recommend this on Google

ILLUSTRATION H.3: Fin de l'article du blog. (source <http://reggies-words.blogspot.com/>)

H.2. Présentation d'un document de notre base

```
{ "GsearchResultClass": "GimageSearch",
  "content": "Omotesando Dori, <b>Tokyo</b>, <b>1993</b>.",
  "contentNoFormatting": "Omotesando Dori, Tokyo, 1993.",
  "height": "891",
  "imageId": "ANd9GcQlhZR1FR8klEuIRbqhaIuS1BNnjRGeCtk115PsvoXs6i_hsed1Wos3HIg",
  "originalContextUrl": "http://reggies-words.blogspot.com/",
  "tbHeight": "146",
  "tbUrl": "http://t3.gstatic.com/images?q=tbn:ANd9GcQlhZR1FR8klEuIRbqhaIuS1
    BNnjRGeCtk115PsvoXs6i_hsed1Wos3HIg",
  "tbWidth": "114",
  "title": "Reggie&#39;s Words",
  "titleNoFormatting": "Reggie&#39;s Words",
  "unescapedUrl": "https://lh6.googleusercontent.com/-PASS832LX_w/TX78narG07I/
    AAAAAAAAAEM/YqF06gSCf84/s1600/scan0001.jpg",
  "url": "https://lh6.googleusercontent.com/-PASS832LX_w/TX78narG07I/
    AAAAAAAAAEM/YqF06gSCf84/s1600/scan0001.jpg",
  "visibleUrl": "reggies-words.blogspot.com",
  "width": "695" }
```

ILLUSTRATION H.4: Exemple de fichier snippet retourné par Google image que nous avons stocké dans des fichiers json. Le contenu du fichier « 0_5945de39-75fe-4570-8f23-2d7a18d5276a.json ».



ILLUSTRATION H.5: Image récupérée associée au contenu du blog.

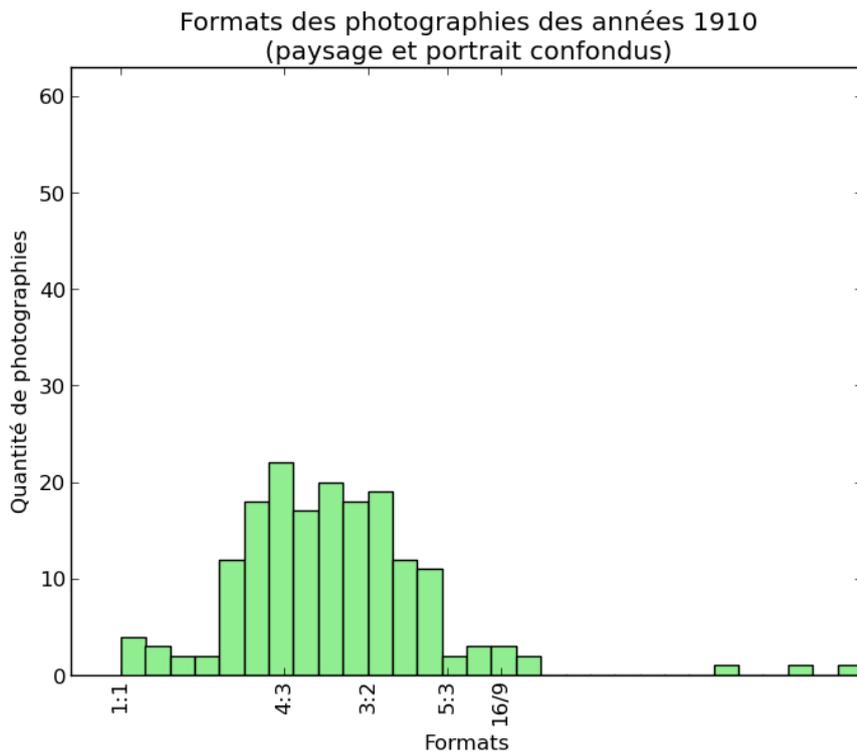


ILLUSTRATION H.7: Distribution du nombre d'images en fonction de leurs formats dans notre jeu de données *web-photo* 1900-2000 pour la décennie 1910.

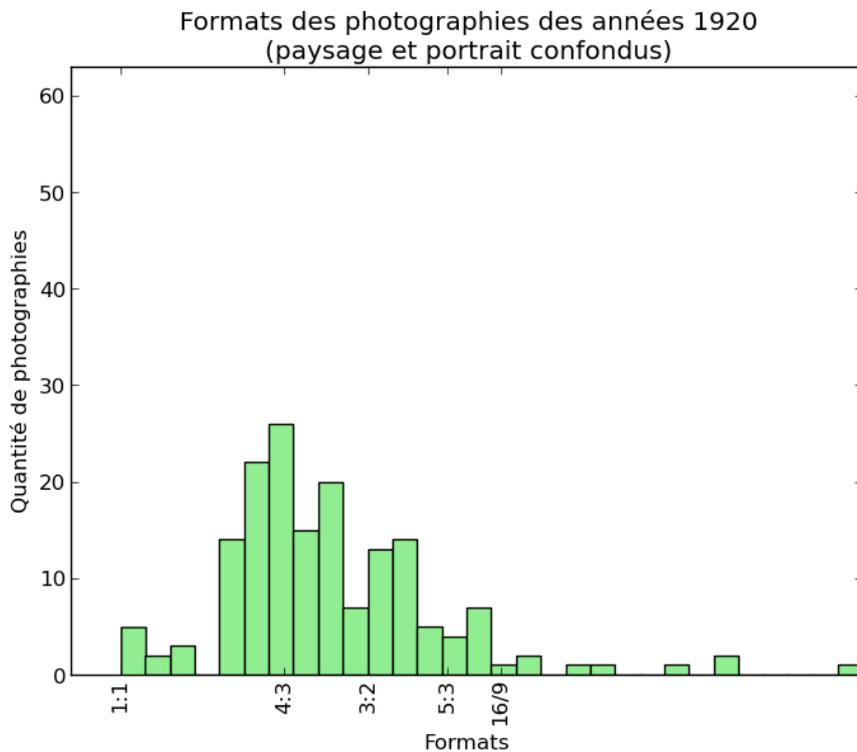


ILLUSTRATION H.8: Distribution du nombre d'images en fonction de leurs formats dans notre jeu de données *web-photo* 1900-2000 pour la décennie 1920.

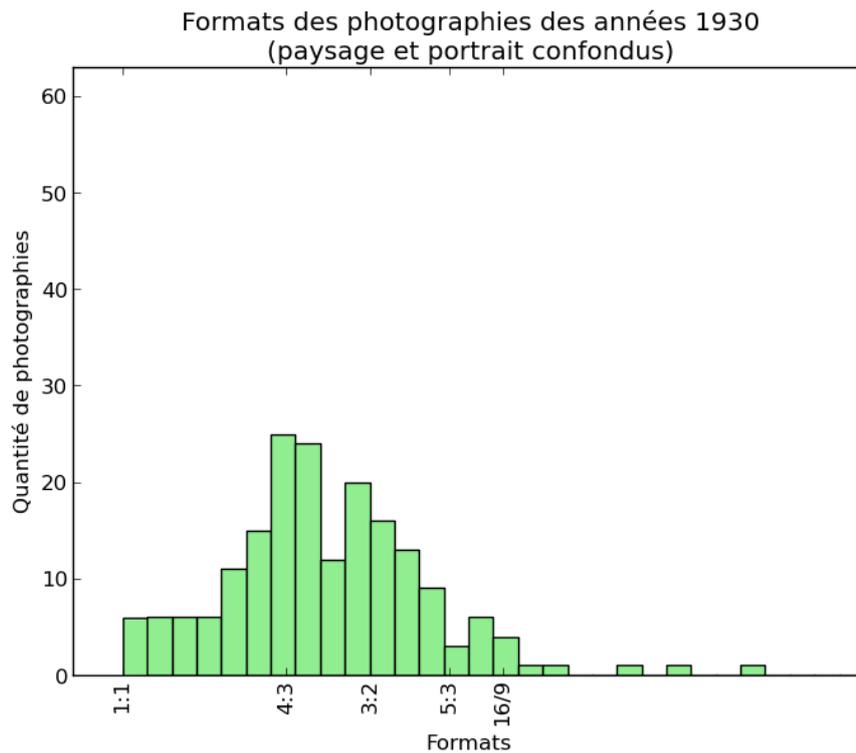


ILLUSTRATION H.9: Distribution du nombre d'images en fonction de leurs formats dans notre jeu de données *web-photo* 1900-2000 pour la décennie 1930.

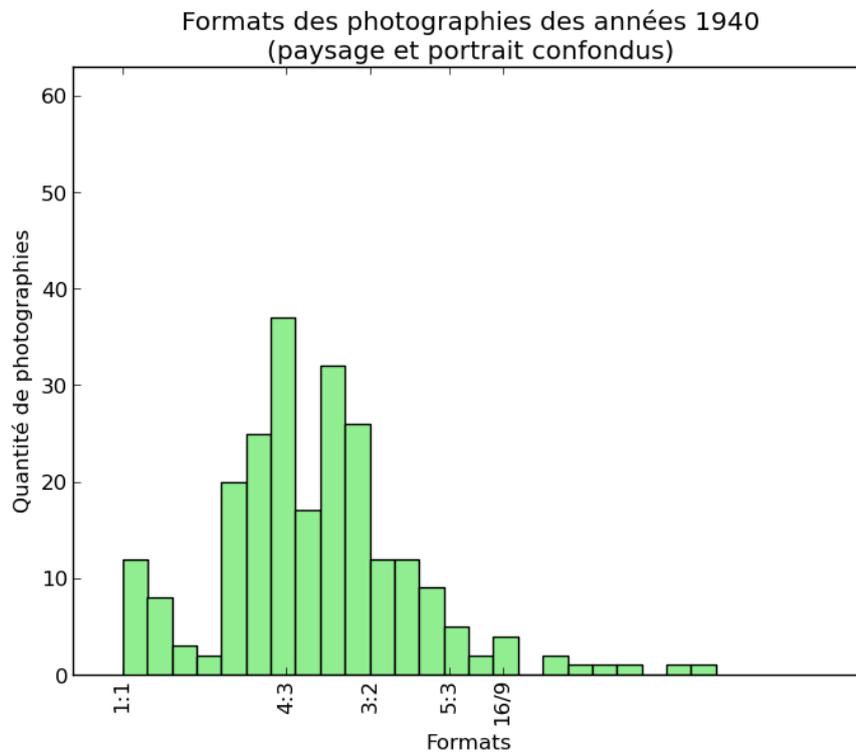


ILLUSTRATION H.10: Distribution du nombre d'images en fonction de leurs formats dans notre jeu de données *web-photo* 1900-2000 pour la décennie 1940.

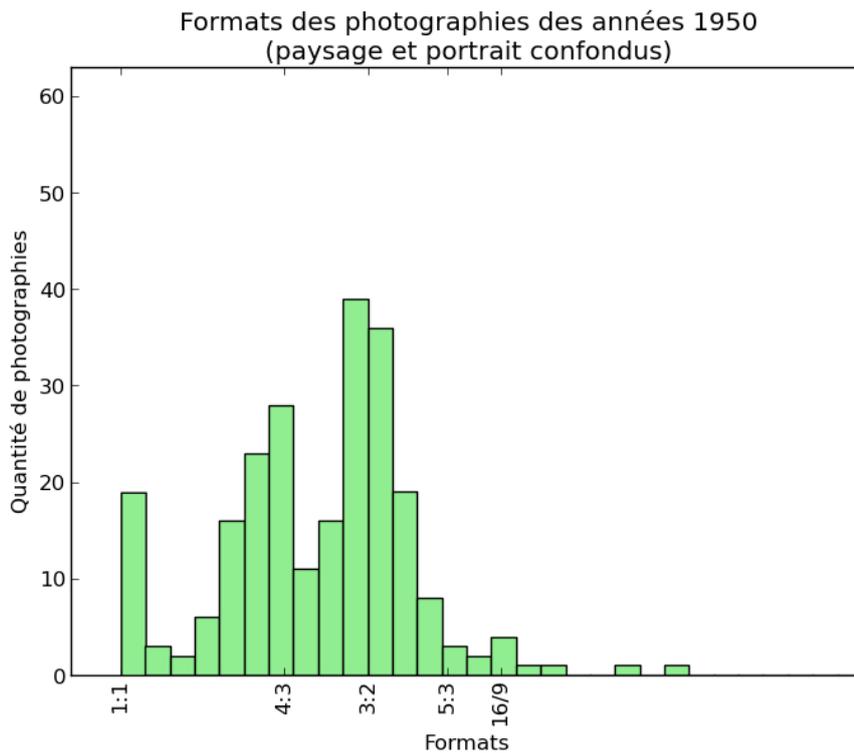


ILLUSTRATION H.11: Distribution du nombre d'images en fonction de leurs formats dans notre jeu de données *web-photo* 1900-2000 pour la décennie 1950.

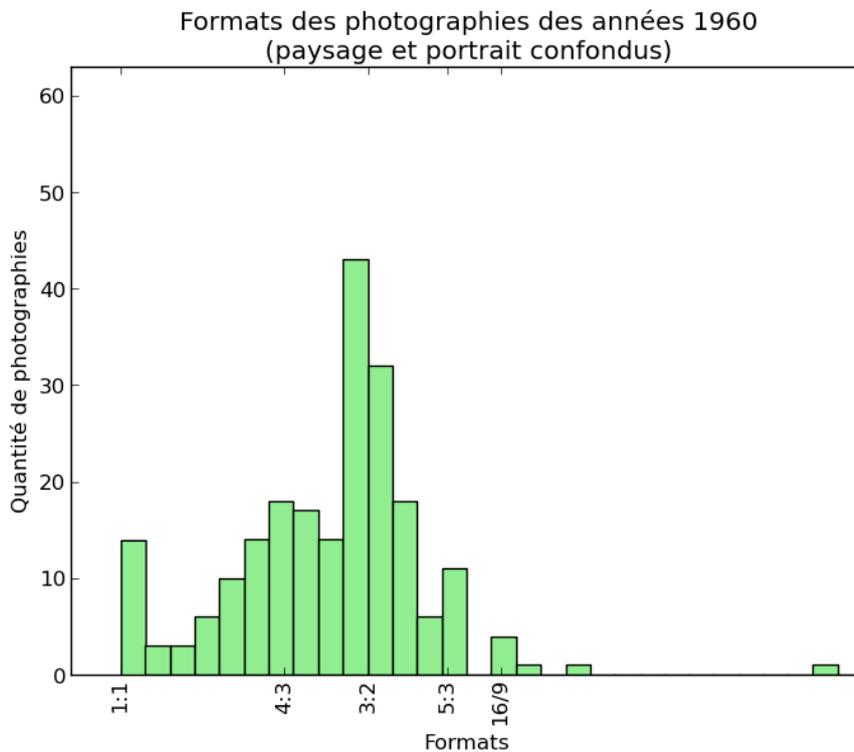


ILLUSTRATION H.12: Distribution du nombre d'images en fonction de leurs formats dans notre jeu de données *web-photo* 1900-2000 pour la décennie 1960.

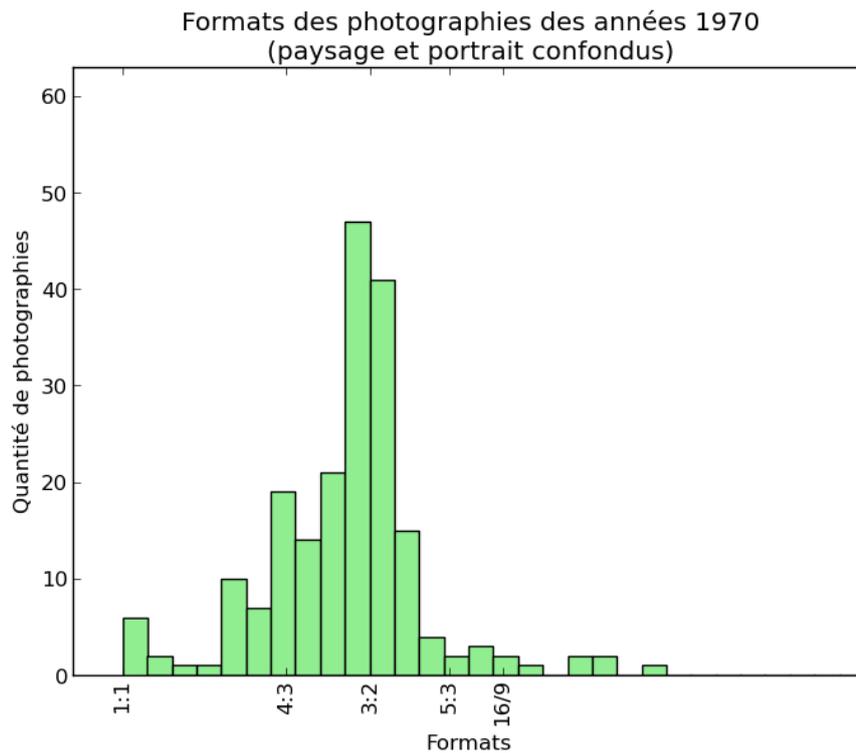


ILLUSTRATION H.13: Distribution du nombre d'images en fonction de leurs formats dans notre jeu de données *web-photo* 1900-2000 pour la décennie 1970.

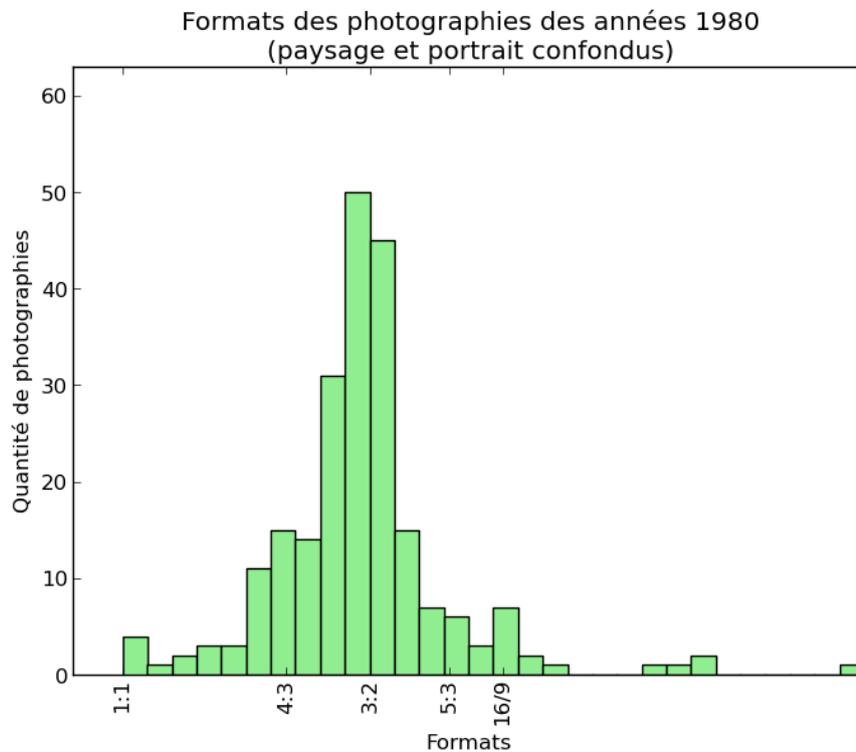


ILLUSTRATION H.14: Distribution du nombre d'images en fonction de leurs formats dans notre jeu de données *web-photo* 1900-2000 pour la décennie 1980.

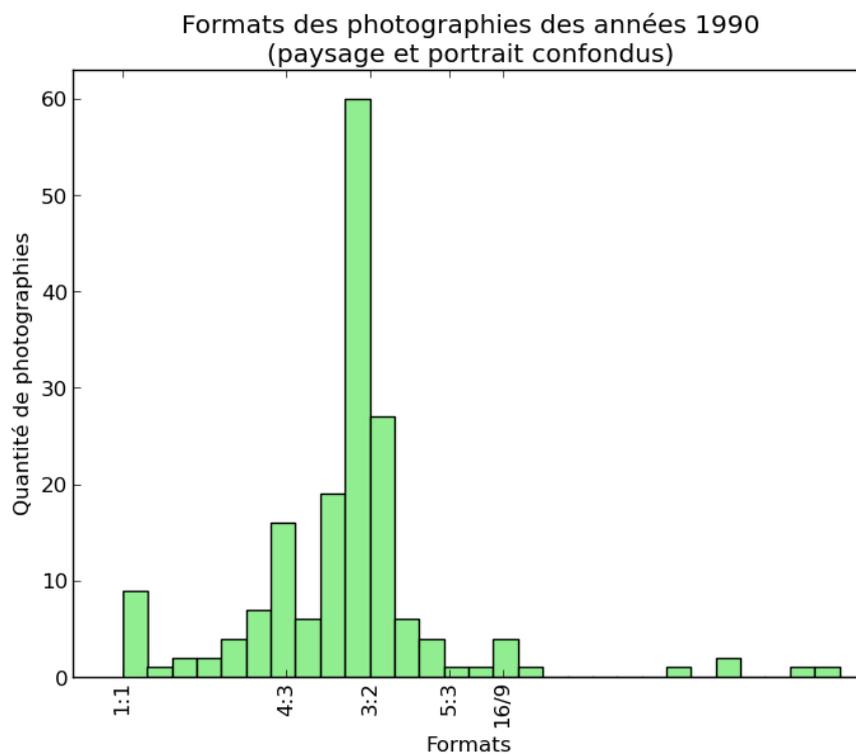


ILLUSTRATION H.15: Distribution du nombre d'images en fonction de leurs formats dans notre jeu de données *web-photo* 1900-2000 pour la décennie 1990.

Annexe I

Observations sur le protocole expérimental

Comme nous l'avons indiqué, dans l'article de PALERMO, HAYS et EFROS (2012) des erreurs typographiques sont présentes. Le jeu de données de PALERMO, HAYS et EFROS (2012) contient en réalité 1 325 images (au lieu de 1 375). Ce qui veut dire que chaque classe possède $1325/5 = 265$ images (au lieu de 275). Pour chacune de ces classes, PALERMO, HAYS et EFROS (2012) ont utilisé aléatoirement 50 images pour le test, ce qui laisse 215 images (au lieu de 225) pour l'apprentissage. Ce qui produit, toutes classes confondues, 250 images pour le test et 1 075 images pour l'apprentissage à chaque pli. Au total sur les 10 plis ce protocole produit 2 500 images pour le test et 10 750 images pour l'apprentissage.

(FERNANDO, MUSELET, KHAN et TUYTELAARS, 2014) ont utilisé par erreur 225 images dans chaque classe pour l'apprentissage. Cette erreur a impacté la taille du jeu de test. Celui-ci ne contient par conséquent que 40 images par classe. Ce qui produit, toutes classes confondues, 200 images pour le test et 1 125 images pour l'apprentissage à chaque pli. Au total sur les 10 plis ce protocole produit 2 000 images pour le test et 11 250 images pour l'apprentissage.

Nous reportons la matrice de confusion indiquée dans l'article de (FERNANDO, MUSELET, KHAN et TUYTELAARS, 2014) dans l'illustration I.1. Nous pouvons observer qu'il manque 10 images pour la deuxième classe. En effet, si nous sommions le nombre d'images de test par colonne nous obtenons ceci :

- 1^{re} colonne : $379 + 9 + 2 + 3 + 7 = 400$ images de test ;
- 2^e colonne : $11 + 326 + 15 + 18 + 20 = 390$ images de test ;
- 3^e colonne : $6 + 24 + 340 + 7 + 23 = 400$ images de test ;
- 4^e colonne : $4 + 7 + 16 + 338 + 35 = 400$ images de test ;
- 5^e colonne : $9 + 25 + 21 + 27 + 318 = 400$ images de test.

Pour un total de 1 990 images de test au lieu de 2 500 selon le protocole expérimental de PALERMO, HAYS et EFROS (2012).

Confusion Matrix

Output Class	1	379 19.0%	11 0.6%	6 0.3%	4 0.2%	9 0.5%	92.7% 7.3%
	2	9 0.5%	326 16.4%	24 1.2%	7 0.4%	25 1.3%	83.4% 16.6%
	3	2 0.1%	15 0.8%	340 17.1%	16 0.8%	21 1.1%	86.3% 13.7%
	4	3 0.2%	18 0.9%	7 0.4%	338 17.0%	27 1.4%	86.0% 14.0%
	5	7 0.4%	20 1.0%	23 1.2%	35 1.8%	318 16.0%	78.9% 21.1%
			94.8% 5.2%	83.6% 16.4%	85.0% 15.0%	84.5% 15.5%	79.5% 20.5%
		1	2	3	4	5	
		Target Class					

ILLUSTRATION I.1: Matrice de confusion de l'article de (FERNANDO, MUSELET, KHAN et TUYTELAARS, 2014).

Annexe J

Comparaison de méthodes de classification ordinaire

Nous reportons ici une série de tableaux comparatifs présentant nos résultats ainsi que ceux obtenus dans l'article de SÁNCHEZ-MONEDERO, GUTIÉRREZ, ANTONIO, TINO et HERVÁS-MARTÍNEZ (2013). Dans chaque tableau comparatif, les résultats reportés sont ceux obtenus avec notre méthode de classification ordinaire présentée dans le chapitre 6 et ceux de SÁNCHEZ-MONEDERO, GUTIÉRREZ, ANTONIO, TINO et HERVÁS-MARTÍNEZ (2013) en terme d' ACC , de MAE , de $AMAE$ et de τ_b . Ces différents résultats sont reportés dans les tableaux J.1, J.2, J.3 et J.4.

Méthodes	Jeux de données									
	automobile	bondrate	contact-lenses	eucalyptus	newthyroid	pasture	squash-stored	squash-unstored	tae	winequality-red
ASAOR(C4.5)	0.696 _{0.059}	0.533 _{0.074}	0.750 _{0.085}	0.639 _{0.036}	0.917 _{0.039}	0.752 _{0.145}	0.603 _{0.118}	0.774 _{0.101}	0.395 _{0.058}	0.603 _{0.021}
GPOR	0.611 _{0.073}	0.578 _{0.032}	0.606 _{0.093}	0.686 _{0.034}	0.966 _{0.024}	0.522 _{0.178}	0.451 _{0.101}	0.644 _{0.162}	0.328 _{0.041}	0.606 _{0.015}
KLDOR	0.722 _{0.058}	0.542 _{0.087}	0.589 _{0.174}	0.611 _{0.028}	0.972 _{0.019}	0.678 _{0.125}	0.703 _{0.112}	0.828 _{0.104}	0.555 _{0.052}	0.603 _{0.017}
POM	0.467 _{0.194}	0.344 _{0.161}	0.622 _{0.138}	0.159 _{0.036}	0.972 _{0.022}	0.496 _{0.154}	0.382 _{0.152}	0.349 _{0.143}	0.512 _{0.089}	0.594 _{0.017}
SVC	0.697 _{0.062}	0.556 _{0.069}	0.794 _{0.129}	0.653 _{0.037}	0.967 _{0.025}	0.633 _{0.134}	0.656 _{0.127}	0.700 _{0.082}	0.539 _{0.062}	0.636 _{0.021}
RED-SVM	0.684 _{0.055}	0.553 _{0.073}	0.700 _{0.111}	0.651 _{0.024}	0.969 _{0.022}	0.648 _{0.134}	0.664 _{0.104}	0.749 _{0.086}	0.522 _{0.074}	0.618 _{0.022}
SVOREX	0.665 _{0.068}	0.553 _{0.096}	0.650 _{0.127}	0.647 _{0.029}	0.967 _{0.022}	0.630 _{0.125}	0.628 _{0.133}	0.718 _{0.128}	0.581 _{0.060}	0.629 _{0.022}
SVORIM	0.639 _{0.076}	0.547 _{0.092}	0.633 _{0.127}	0.639 _{0.028}	0.969 _{0.021}	0.667 _{0.120}	0.639 _{0.118}	0.764 _{0.103}	0.590 _{0.066}	0.630 _{0.022}
SVR-PCDOC	0.678 _{0.060}	0.540 _{0.101}	0.689 _{0.095}	0.648 _{0.029}	0.973 _{0.020}	0.656 _{0.103}	0.685 _{0.123}	0.695 _{0.084}	0.582 _{0.064}	0.631 _{0.022}
Notre approche	0.619 _{0.071}	0.518 _{0.086}	0.733 _{0.126}	0.639 _{0.031}	0.969 _{0.022}	0.626 _{0.142}	0.603 _{0.111}	0.733 _{0.103}	0.454 _{0.079}	0.589 _{0.016}

TABLEAU J.1: Comparatif des résultats de notre méthode de classification ordinale présentée dans le chapitre 6 avec ceux de SÁNCHEZ-MONEDERO, GUTIÉRREZ, ANTONIO, TINO et HERVÁS-MARTÍNEZ (2013) en terme d' $ACC_{\text{écart type}}$.

Méthodes	Jeux de données									
	automobile	bondrate	contact-lenses	eucalyptus	newthyroid	pasture	squash-stored	squash-unstored	tae	winequality-red
ASAOR(C4.5)	0.401 _{0.095}	0.624 _{0.079}	0.367 _{0.154}	0.384 _{0.042}	0.083 _{0.039}	0.248 _{0.145}	0.444 _{0.140}	0.239 _{0.109}	0.686 _{0.146}	0.441 _{0.023}
GPOR	0.594 _{0.131}	0.624 _{0.062}	0.511 _{0.175}	0.331 _{0.038}	0.034 _{0.024}	0.489 _{0.190}	0.626 _{0.148}	0.356 _{0.162}	0.861 _{0.155}	0.425 _{0.017}
KLDOR	0.334 _{0.076}	0.587 _{0.107}	0.539 _{0.208}	0.424 _{0.032}	0.028 _{0.019}	0.322 _{0.125}	0.308 _{0.128}	0.172 _{0.104}	0.473 _{0.069}	0.443 _{0.019}
POM	0.953 _{0.687}	0.947 _{0.321}	0.533 _{0.241}	2.029 _{0.070}	0.028 _{0.022}	0.585 _{0.204}	0.813 _{0.248}	0.826 _{0.230}	0.626 _{0.126}	0.439 _{0.019}
SVC	0.446 _{0.095}	0.624 _{0.090}	0.311 _{0.222}	0.394 _{0.042}	0.033 _{0.025}	0.367 _{0.134}	0.377 _{0.160}	0.308 _{0.090}	0.578 _{0.083}	0.408 _{0.020}
RED-SVM	0.393 _{0.073}	0.598 _{0.088}	0.378 _{0.169}	0.380 _{0.027}	0.032 _{0.022}	0.359 _{0.142}	0.346 _{0.110}	0.251 _{0.086}	0.515 _{0.087}	0.419 _{0.021}
SVOREX	0.408 _{0.089}	0.573 _{0.121}	0.489 _{0.185}	0.392 _{0.031}	0.033 _{0.022}	0.370 _{0.125}	0.382 _{0.139}	0.282 _{0.128}	0.485 _{0.078}	0.408 _{0.023}
SVORIM	0.424 _{0.090}	0.591 _{0.102}	0.506 _{0.167}	0.395 _{0.035}	0.031 _{0.021}	0.333 _{0.120}	0.372 _{0.126}	0.239 _{0.109}	0.461 _{0.081}	0.406 _{0.022}
SVR-PCDOC	0.397 _{0.093}	0.568 _{0.126}	0.367 _{0.154}	0.392 _{0.038}	0.027 _{0.020}	0.348 _{0.104}	0.326 _{0.141}	0.305 _{0.084}	0.457 _{0.071}	0.400 _{0.023}
Notre approche	0.501 _{0.117}	0.593 _{0.112}	0.383 _{0.173}	0.403 _{0.036}	0.031 _{0.022}	0.378 _{0.148}	0.444 _{0.128}	0.267 _{0.103}	0.623 _{0.066}	0.449 _{0.016}

TABLEAU J.2: Comparatif des résultats de notre méthode de classification ordinale présentée dans le chapitre 6 avec ceux de SÁNCHEZ-MONEDERO, GUTIÉRREZ, ANTONIO, TINO et HERVÁS-MARTÍNEZ (2013) en terme de $MAE_{\text{écart type}}$.

Méthodes	Jeux de données									
	automobile	bondrate	contact-lenses	eucalyptus	newthyroid	pasture	squash-stored	squash-unstored	tae	winequality-red
ASAOR(C4.5)	0.511 _{0.104}	1.226 _{0.175}	0.315 _{0.124}	0.428 _{0.045}	0.115 _{0.056}	0.248 _{0.145}	0.502 _{0.192}	0.256 _{0.149}	0.689 _{0.151}	1.045 _{0.080}
GPOR	0.792 _{0.200}	1.360 _{0.122}	0.651 _{0.286}	0.362 _{0.040}	0.062 _{0.049}	0.489 _{0.190}	0.797 _{0.234}	0.443 _{0.226}	0.863 _{0.164}	1.065 _{0.065}
KLDOR	0.345 _{0.104}	1.037 _{0.270}	0.519 _{0.280}	0.426 _{0.038}	0.059 _{0.040}	0.322 _{0.125}	0.349 _{0.156}	0.309 _{0.180}	0.471 _{0.070}	1.258 _{0.069}
POM	1.026 _{0.800}	1.103 _{0.403}	0.535 _{0.275}	1.990 _{0.048}	0.050 _{0.040}	0.585 _{0.204}	0.815 _{0.251}	0.791 _{0.332}	0.627 _{0.128}	1.085 _{0.037}
SVC	0.486 _{0.125}	1.265 _{0.183}	0.307 _{0.277}	0.433 _{0.048}	0.060 _{0.051}	0.367 _{0.134}	0.446 _{0.189}	0.444 _{0.163}	0.576 _{0.083}	1.119 _{0.069}
RED-SVM	0.468 _{0.096}	1.184 _{0.225}	0.385 _{0.198}	0.414 _{0.030}	0.057 _{0.049}	0.359 _{0.142}	0.391 _{0.149}	0.348 _{0.159}	0.513 _{0.086}	1.068 _{0.069}
SVOREX	0.518 _{0.096}	1.072 _{0.217}	0.517 _{0.303}	0.411 _{0.034}	0.054 _{0.042}	0.370 _{0.125}	0.433 _{0.172}	0.426 _{0.157}	0.484 _{0.079}	1.095 _{0.067}
SVORIM	0.523 _{0.105}	1.114 _{0.233}	0.589 _{0.259}	0.420 _{0.043}	0.055 _{0.042}	0.333 _{0.120}	0.427 _{0.148}	0.367 _{0.140}	0.459 _{0.081}	1.093 _{0.072}
SVR-PCDOC	0.440 _{0.128}	0.969 _{0.224}	0.420 _{0.098}	0.400 _{0.043}	0.045 _{0.040}	0.348 _{0.104}	0.360 _{0.184}	0.396 _{0.158}	0.455 _{0.071}	1.040 _{0.096}
Notre approche	0.583 _{0.108}	1.024 _{0.165}	0.338 _{0.202}	0.436 _{0.043}	0.049 _{0.045}	0.378 _{0.148}	0.498 _{0.162}	0.359 _{0.157}	0.620 _{0.065}	1.122 _{0.065}

TABLEAU J.3: Comparatif des résultats de notre méthode de classification ordinale présentée dans le chapitre 6 avec ceux de SÁNCHEZ-MONEDERO, GUTIÉRREZ, ANTONIO, TINO et HERVÁS-MARTÍNEZ (2013) en terme de $AMAE_{\text{écart type}}$.

Méthodes	Jeux de données									
	automobile	bondrate	contact-lenses	eucalyptus	newthyroid	pasture	squash-stored	squash-unstored	tae	winequality-red
ASAOR(C4.5)	0.741 _{0.069}	0.143 _{0.159}	0.604 _{0.216}	0.802 _{0.025}	0.853 _{0.067}	0.778 _{0.132}	0.415 _{0.245}	0.692 _{0.145}	0.243 _{0.177}	0.496 _{0.036}
GPOR	0.557 _{0.118}	0.000 _{0.000}	0.348 _{0.304}	0.830 _{0.020}	0.938 _{0.045}	0.461 _{0.314}	0.075 _{0.211}	0.420 _{0.331}	-0.018 _{0.108}	0.523 _{0.026}
KLDOR	0.793 _{0.056}	0.356 _{0.257}	0.448 _{0.273}	0.786 _{0.017}	0.948 _{0.034}	0.718 _{0.133}	0.646 _{0.160}	0.764 _{0.161}	0.477 _{0.114}	0.460 _{0.028}
POM	0.495 _{0.283}	0.290 _{0.302}	0.458 _{0.309}	0.008 _{0.038}	0.949 _{0.040}	0.463 _{0.237}	0.169 _{0.304}	0.109 _{0.305}	0.317 _{0.129}	0.497 _{0.025}
SVC	0.695 _{0.077}	0.121 _{0.177}	0.601 _{0.300}	0.783 _{0.025}	0.939 _{0.045}	0.698 _{0.133}	0.541 _{0.240}	0.599 _{0.140}	0.375 _{0.110}	0.516 _{0.027}
RED-SVM	0.751 _{0.054}	0.254 _{0.247}	0.577 _{0.242}	0.800 _{0.017}	0.943 _{0.041}	0.707 _{0.129}	0.601 _{0.148}	0.662 _{0.108}	0.417 _{0.120}	0.525 _{0.030}
SVOREX	0.749 _{0.062}	0.369 _{0.216}	0.425 _{0.304}	0.794 _{0.019}	0.941 _{0.040}	0.691 _{0.115}	0.534 _{0.207}	0.592 _{0.212}	0.445 _{0.110}	0.531 _{0.028}
SVORIM	0.748 _{0.065}	0.299 _{0.230}	0.382 _{0.269}	0.792 _{0.020}	0.944 _{0.038}	0.710 _{0.114}	0.542 _{0.167}	0.656 _{0.187}	0.482 _{0.118}	0.533 _{0.030}
SVR-PCDOC	0.744 _{0.076}	0.455 _{0.218}	0.620 _{0.217}	0.795 _{0.024}	0.952 _{0.037}	0.712 _{0.102}	0.610 _{0.201}	0.602 _{0.133}	0.493 _{0.101}	0.542 _{0.033}
Notre approche	0.688 _{0.078}	0.435 _{0.135}	0.557 _{0.252}	0.789 _{0.020}	0.945 _{0.039}	0.698 _{0.138}	0.486 _{0.171}	0.648 _{0.134}	0.286 _{0.111}	0.492 _{0.022}

TABLEAU J.4: Comparatif des résultats de notre méthode de classification ordinale présentée dans le chapitre 6 avec ceux de SÁNCHEZ-MONEDERO, GUTIÉRREZ, ANTONIO, TINO et HERVÁS-MARTÍNEZ (2013) en terme de τ_b écart type.

Datation automatique de photographies à partir de caractéristiques textuelles et visuelles

Dans cette thèse, la problématique que nous traitons est celle de la datation automatique de photographies anciennes à partir d'attributs issus de l'image et du texte qui l'entoure.

Au cours de ces derniers siècles, la photographie et les métadonnées ont subi de nombreuses mutations. Dans une vocation informative et contextualisante, nous reportons nos observations au regard de ces avancées technologiques qui permettent de comprendre l'évolution des méthodes de datation de photographies, historiquement manuelles et récemment devenues automatiques.

Nous nous sommes intéressés à une notion fondamentale et intrinsèque de la datation : le caractère ordinal du temps. Nous avons donc étudié les méthodes de classification ordinale et les mesures d'évaluation associées.

Pour cette thèse, nous avons construit une large collection de pages web associant clichés anciens et contenus textuels. Celle-ci a pour vocation de permettre l'entraînement et l'évaluation des modèles proposés. Ceci représente une avancée dans ce domaine car aucun jeu de données comparable n'existait auparavant.

Nous proposons un modèle original que nous avons développé et évalué sur des bases de photographies pour les dater grâce à leur contenu visuel. La particularité majeure de notre approche est l'exploitation pour la première fois du caractère ordinal du temps.

Enfin, nous explorons la datation de photographies par l'usage des données textuelles environnantes. Notre démarche imite le comportement humain utilisé lors de la datation de photographies en contexte. Ainsi nous proposons pour la première fois une méthode automatique permettant de dater des clichés par l'usage de l'environnement textuel.

Mots-clés : Photographies - - Datation ; Classification automatique ; Recherche de l'information ; Vision par ordinateur ; Traitement automatique du langage naturel ; Entités nommées.

Automatic photographs dating from textual and visual features

In this thesis, we address the problem of automatic dating of old photographs using attributes from image and surrounding text.

Over the last centuries, photography and metadata have undergone many mutations. In an informative and contextualizing vocation, we report our observations regarding these technological advances. This information helps to understand the evolution of photographs dating methods that have historically been studied manually before becoming recently automatic.

We focused in a fundamental concept of dating : the ordinal nature of time. To this extent we studied ordinal classification methods and associated evaluation measures.

In this thesis, we have built a large collection of web pages combining old photographs and text content. This collection aims at allowing both training and evaluation of the proposed models. This represents an advance in the field of photographs dating because no comparable data set previously existed.

We propose an original model that we have developed and evaluated on image databases to date the photographs with their visual content. The major feature of our approach is the exploitation of the ordinal structure of time. This is to our knowledge the first use of this intrinsic characteristic of photographs dating.

Moreover, we explore the dating of photographs by the use of the surrounding text data. In our approach, we mimic the human behavior used in the context of photographs dating. Thus we propose for the first time an automatic method for the dating of photographs by the use of the textual environment.

Keywords : Photographs - - Dating ; Automatic classification ; Information retrieval ; Computer vision ; Natural language processing (Computer science) ; Named entities.

Discipline : Informatique et applications

Laboratoire : GREYC (UMR 6072) – Université de Caen Normandie
Campus Côte de Nacre, Boulevard du Maréchal Juin 14 032 CAEN cedex 5