Chapter 1

Literature search

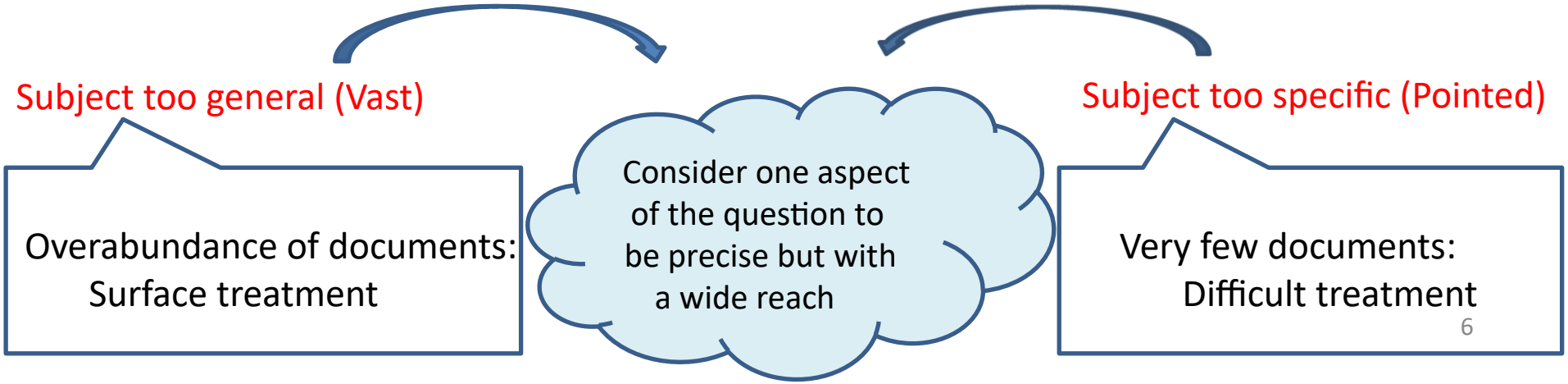
Literature review: Introduction

The initial stage in carrying out an EFP is documentary research. This consists of searching for and identifying documents relating to the topic under study. However, the overabundance of information available on the Web requires using a good research strategy to establish a compromise between the

wealth of documentation and the scientific rigor expressing the quality and reliability of the information. To do this, you need to select the information sources you will use. However, before starting to search the literature, you must have a subject!

Documentary research: Definition of the subject

The definition of a subject for an EFP is done either through an internship in a socio-economic institution or by the teacher supervising you. So, the first stage in the search for information is to consult documents that will help clarify the subject and give you a general understanding of it. There are two possible difficulties here:



Documentary research: Sources of information

Once you have defined your subject, you need to choose the best sources of information for your documentary research. Generally, for a PFE, there are four primary sources of information:

- Conference papers or publications in internationally renowned journals,
- Books,
- Theses (Masters or Doctorate),
- Website,

✞ Articles:

Documentary research: Sources of information

Whatever the subject, the number of articles available on the web is vast. So, to ensure the reliability of the information, any article whose source is unknown should be discarded.

So where are the reliable journals with good-quality articles?

The journal and the conference must be sponsored by a well-known publisher such as IEEE, ELSEVIER, SPRINGER, ACM, etc.

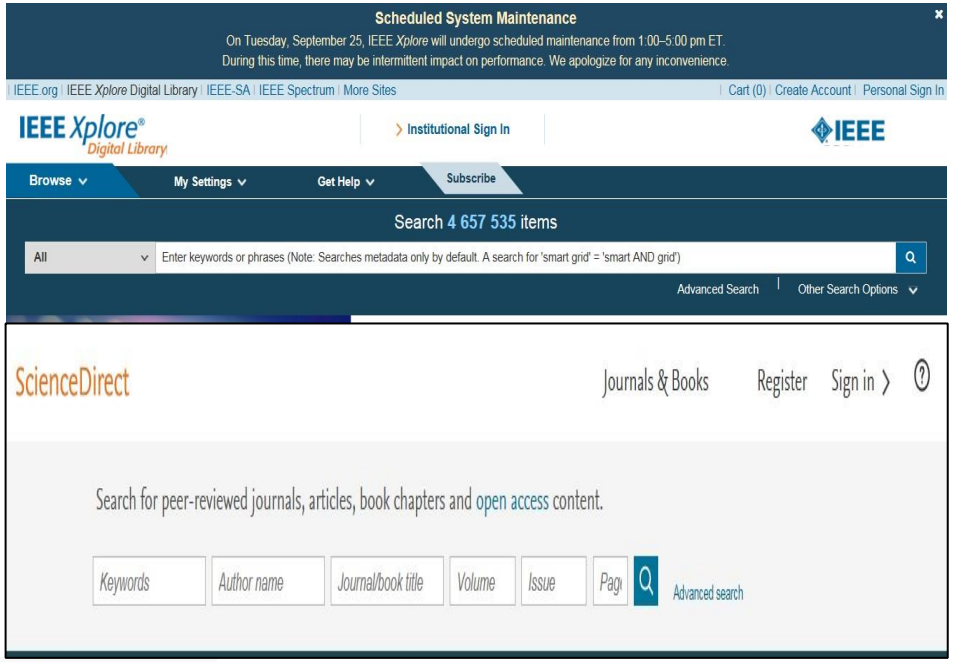
It should be noted that the first three publishers are the primary references for all areas of research addressed by electronics specialists.

Documentary research: Sources of information

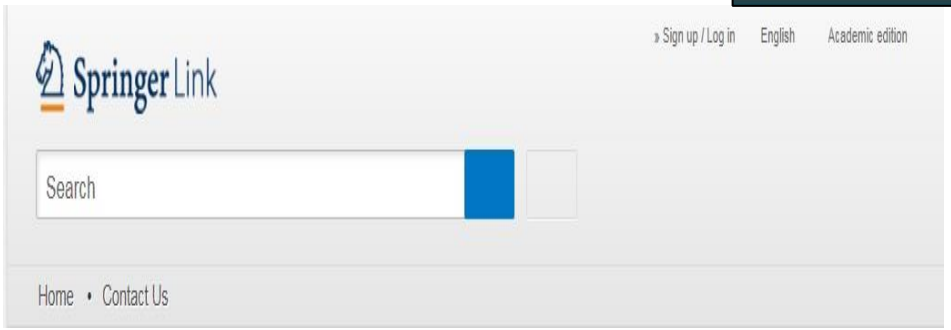
- Institute of Electrical and Electronics Engineers: IEEE
- ELSEVIER Company
- SPRINGER SCIENCE + BUSINESS MEDIA
- Association for Computer Machinery: ACM
- Institution of Engineering and Technology: IET

Documentary research: Sources of information

Looking at the publishers we have just mentioned, we can see that each has its bookshop:



- IEEE: IEEE Xplore
- ELSEVIER: Science Direct
- SPRINGER: Springer Link



Documentary research: Sources of information

ELSEVIER publication

Expert Systems With Applications 51 (2016) 186–194

Contents lists available at ScienceDirect

Expert Systems With Applications

journal homepage: www.elsevier.com/locate/eswa

New off-line Handwritten Signature Verification method based on Artificial Immune Recognition System

Yasmine Serdouk*, Hassiba Nemmour, Youcef Chibani

Communicating and Intelligent System Engineering Laboratory, Faculty of Electronic and Computer Sciences, University of Sciences and Technology Houari Boumediene (USTHB), Bab-Ezzouar El-Alia BP 32, 16111 Algiers, Algeria

ARTICLE INFO

Keywords:
Artificial Immune Recognition System
Gradient Local Binary Patterns
Longest Run Features
Signature verification

ABSTRACT

Natural Immune System offers many interesting features that inspired the design of Artificial Immune Systems (AIS) used to solve various problems of engineering and artificial intelligence. AIS are particularly successful in fault detection and diagnosis applications where anomalies such as errors and failures are assimilated to viruses that should be detected. Thereby, AIS seem suitable to automatically detect forgeries in signature verification systems. This paper proposes a novel method for off-line signature verification that is based on the Artificial Immune Recognition System (AIRS). For feature generation, two different descriptors are proposed to generate signature traits. The first is the Gradient Local Binary Patterns that estimates gradient features based on the LBP neighborhood. The second descriptor is the Longest Run Feature, which describes the signature topology by considering longest suites of text pixels. Performance evaluation is carried out on CEDAR and GPDS-100 datasets. The results obtained showed that the proposed system has promising performance and often comfortably outperforms the state of the art.

1. Introduction

Biometrics is the science which identifies persons on the basis of their biological or behavioral features. Currently, various identification systems employing fingerprints, hand geometry, retina, iris, or faces provide very interesting performances. Nevertheless, in some specific applications using paper documents such as bank checks and contracts, handwritten signature remains the oldest and the main identification tool. Many successful studies have been reported on this subject. However, because of speed and robustness requirements, signature verification remains an open research issue (Impedovo, Modugno, Pirlo, & Stasolla, 2008). There are two approaches for developing such systems, the on-line verification and the off-line verification (Impedovo et al., 2008; Plamondon & Srihari, 2000). In the former, signatures are acquired via an electronic device such as tablets and an instrumented stylus that captures temporary information such as the x–y coordinates, velocity or acceleration (Plamondon & Srihari, 2000). In the off-line verification, such information cannot be recovered since signatures are prewritten on a paper. Consequently, it is less accurate but it has more practical application areas. On the other hand, the verification itself may be attempted in either writer-dependent or writer-independent context. In the writer-dependent case, the system is trained by using genuine signatures or both genuine and forgeries of a specific writer (Rivar, Granger, & Sabourin, 2013). This means that the training process should be repeated each time a new signer is presented to the system. The writer-independent scenario develops a generic system that can be tested on any new writer (Kumar, Sharma, & Chanda, 2012). Precisely, the system is developed using dissimilarities between genuine and forged signatures of some writers and tested on dissimilarities of other writers.

Note that since signatures are strong variable entities, even for human experts, their verification is not a trivial matter (Vargas, Ferrer, Travieso, & Alonso, 2011). For this reason, the signature verification has been widely investigated during the past years. The literature reports a lot of research works using various feature generation and verification methods (Impedovo et al., 2008; Plamondon & Srihari, 2000). Presently, we are focused on the writer-dependent off-line signature verification. In this approach, the best off-line signature verifiers (when tested on publicly available databases and against skilled forgeries) deliver error rates of approximately 9–10% (Kovari & Charaf, 2013). Such scores reveal that more theoretical advances should be done in the signature verification field in order to reach real-life verification systems.

* Corresponding author. Tel.: +213 561122950.
E-mail addresses: yserdouk@usthb.dz, serdoukyasmine@gmail.com (Y. Serdouk), hnemmour@usthb.dz (H. Nemmour), ychibani@usthb.dz (Y. Chibani).

http://dx.doi.org/10.1016/j.eswa.2016.01.001
0957-4174/

Documentary research: Sources of information

SPRINGER publication

Pattern Analysis and Applications
https://doi.org/10.1007/s10044-018-0735-y

THEORETICAL ADVANCES

Hybrid one-class classifier ensemble based on fuzzy integral for open-lexicon handwritten Arabic word recognition

Bilal Hadjadji[1] · Youcef Chibani[1] · Hassiba Nemmour[1]

Received: 31 January 2017 / Accepted: 19 July 2018

Abstract
One-class classifier (OCC) is involved for solving different kinds of problems due to its ability to represent a class distribution regardless the remaining classes. Its main advantage for multi-class classification is offering an open system and therefore allows easily extending new classes without retraining OCCs. So far, hidden Markov models, support vector machines and neural networks are the most used classifiers for Arabic word recognition, which provides a system with closed lexicon. In this paper, the OCCs are explored in order to perform an Arabic word recognition system with an open lexicon. Generally, pattern recognition systems designed by a single system suffer from limitations such as the lack of uniqueness and non-universality. Thus, combining multiple systems becomes an attractive research topic for performance and robustness enhancement. Fixed rules are commonly used as combiners for the hybrid OCC ensembles. The present paper aims to propose a combination scheme of OCCs based on the use of fuzzy integral (FI) operators. Furthermore, an alternative framework is proposed to design a parameter-independent and open-lexicon handwritten Arabic word recognition system as well as a new density measure function. Experimental results conducted on Arabic handwritten dataset using different types of OCCs with large number of classes highlight the superiority of FI for hybrid OCC ensembles.

Keywords One-class classifiers · Hybrid OCC ensemble · Fuzzy integral · Density measures · Open-lexicon Arabic word recognition

1 Introduction

Handwritten Arabic word recognition is an active research field due to its interesting use in different applications such as automatic sorting of postal mail, automatic bank check processing, bills processing, passport validation and, recently, for historical document reading via text to speech applications, helping blinds to read and recognizing handwritten historical documents [1–5]. Unlike Latin languages, Arabic is written from right to left. Also, it has its own diacritical marking such as dumma (ˇ), hamza (ء) and the madda (–). Regarding character shapes, Arabic script has two main properties. On the one hand, several letters share the same shape and differ only in the number and position of dots, such as "djim: ج," "ha: ح" and "kha: خ." On the other hand, some letters change their shape according to their position at the beginning, the medial or the end in the word. For instance, the letter "Aïn" can be written through four shapes that are: "ع, ـعـ, ـع, ع," where the two last shapes are related to end positions which change if the word is fully connected or not.

So far, the Arabic word recognition is considered as one of the most challenging tasks of pattern recognition for its specific writing as well as its variability. In this context, the analytical and holistic approaches are the two possible ways for recognizing an Arabic word [6]. The first one consists of segmenting a word image into subwords or isolated characters, which are recognized through character recognition. Generally, this approach is employed when a very large vocabulary is available since it is impossible to construct a

✉ Bilal Hadjadji
bhadjadji@usthb.dz

Youcef Chibani
ychibani@usthb.dz

Hassiba Nemmour
hnemmour@usthb.dz

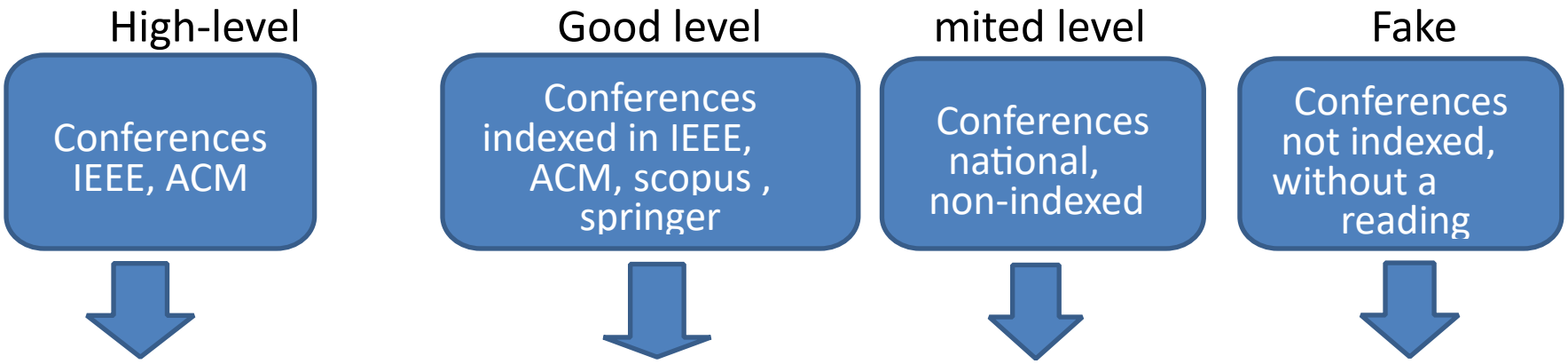
[1] Laboratoire d'Ingénierie des Systèmes Intelligents et Communicants (LISIC), Faculty of Electronics and Computer Science, University of Science and Technology Houari Boumediene (USTHB), 32, El Alia, Bab Ezzouar, 16111 Algiers, Algeria

Published online: 31 July 2018

Documentary research: Sources of information

- Conferences :

Conferences are classified according to the level of papers presented.



The title of the conference begins with: IEEE International conference..., or ACM international conference,...;

Conference with a good review process

To do as first step in research

Conferences ineligible, see CSF website

Documentary research: Objective

In addition to understanding the fundamental concepts of the subject, documentary research helps to understand the aspect being studied. To produce a dissertation or, more specifically, a thesis, you must assess all the work you have already done on the subject. This makes it possible to establish what is known as a STATE OF THE ART (State of the art, background, related works, etc.). The state-of-the-art consists of listing the various jobs carried out on the theme, illustrating the advantages and, above all, the disadvantages, from which issues or questions still to be answered arise.

Therefore, the work of a PFE or a thesis must address a problem or question that is still being asked to propose a solution.

Documentary research: Conclusion

Once the documents have been collected, they must be read before starting the EFP work. From this reading, we will learn the general concepts related to the theme and will also be able to situate the work of the EFP. In this way, the objective (the contribution of a thesis) will be highlighted. Once this stage has been completed, you can begin the work proposed in the EFP.