Knowledge-Based Systems 101 (2016) 60–70

Contents lists available at ScienceDirect

Knowledge-Based Systems

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

Identification of mammography anomalies for breast cancer detection by an ensemble of classification models based on artificial immune system

Gabriele Magna, Paola Casti, Sowmya Velappa Jayaraman, Marcello Salmeri, Arianna Mencattini, Eugenio Martinelli*, Corrado Di Natale

Department of Electronic Engineering, University of Rome "Tor Vergata", Via del Politecnico 1, 00133 Roma, Italy

ARTICLE INFO

Article history:
Received 14 May 2015
Revised 10 February 2016
Accepted 19 February 2016
Available online 28 March 2016

Keywords:
Artificial immune system
Ensemble of classifiers
Breast cancer

ABSTRACT

The interpretation of diagnostic images is often conditioned by the specific properties of the instrument that generated the image. This makes particularly complicated to develop universal recognition algorithms that can facilitate the diagnosis in case of massive population screenings. Mammography is a typical example where such an algorithm is required. Although the technological advances in medical imaging increases the accuracy of interpretation of images, the improved resolution may not facilitate the identification of breast cancer at a very early stage due to the many confounding factors related, for instance, to differences in instrument settings or breast positioning by the operator. Being impossible an exact standardization, the problem of reducing the effects of different instruments and operators can be faced with a proper algorithm that can extract from each image the relevant information in an unsupervised manner, thus limiting the influence of instrumental and positioning issues. For this scope, in this paper we investigated the properties of a classifier based on an ensemble of Adaptive Artificial Immune Networks (A^2INET) applied to original mammography image indicators aimed at diagnosing bilateral asymmetry that is known to be correlated with increased breast cancer risk. Classification models were trained using a set of descriptors measuring the degree of similarity of paired regions of the left and right breasts. Noteworthy, the ensemble of A^2INET models achieved very high classification rates even when training and testing were made on two completely independent and heterogeneous datasets. The obtained results are promising (maximum accuracy level of 0.90, sensitivity of 0.93 and specificity of 0.87) and they prefigure to apply automatic diagnostic tools in clinical practice exploiting a network of different instrument databases.

1. Introduction

For almost fifty years, mainstream artificial intelligence has been gaining popularity in creating computers and algorithms that emulate human cognitive abilities limited to the execution of tasks in well-defined contexts not requiring fundamental aspects of intelligence such as behavioral autonomy, social interaction, evolution, and learning [1]. The mentioned aspects are fundamental to allow organisms to survive in evolving environments. The mimicking of these aspects became popular in the mid-1980 when paradigms of artificial intelligence were introduced shifting the focus to biological processes such as self-organization. Examples of the introduced concepts include evolutionary computation and electronics, artificial neural networks, artificial immune systems, bio-robotics, and swarm intelligence. Among them, the principle of the immune system has been a source of inspiration to design algorithms able to solve complex computational problems [2]. The highly distributed, adaptive, and self-organizing nature, combined with learning, memory, and pattern recognition features gave rise the so called Artificial Immune System (AIS). Such a cooperative effort of immunology and engineering mimics the defense mechanisms used against pathogens and to react to internal faults. AIS has been largely applied to pattern recognition problems [3], noise reduction [4], function optimization [5], synthetic aperture radar imaging [6], and biological modeling, with particular interests to areas of research such as negative selection algorithms [7], clonal selection algorithms [8–10], danger theory and dendritic

* Corresponding author. Tel.: +39 06 72597259; fax: +39 06 2020 519.
E-mail address: martinelli@ing.uniroma2.it (E. Martinelli).

http://dx.doi.org/10.1016/j.knosys.2016.02.019
0950-7051/

cell algorithms [11,12], and the most popular Artificial Immune NETwork (AINET) [13].

Obviously, only the basic principles of the immune systems are actually considered in the AIS algorithms. In order to describe the algorithm, here a brief description of the principle is given.

The immune system reacts to the introduction into the organism of a foreign agent, such as a virus. As a consequence of this event, special cells, the B-cells, begin to produce particular antibodies for each antigen characterizing the foreign agent. When it is sufficiently stimulated thanks to the close affinity between the B-cell and the antigen, the B-cells rapidly produce clones of themselves. Cloned B-cells can also mutate in order to increase the affinity with the antigen. The whole cloning/mutation procedure is repeated many times for each antigen. The defense process continues during time, leaving a storage of a small number of memory-B-cells in order to protect the body from future invasion of pathogens characterized by the same antigens.

Some of the immune system properties can inspire a learning pattern method. First of all, the concept of affinity-based defense states that the B-cells generate antibodies as affine as possible to the antigens. This concept can be imported in a distance-based pattern recognition systems. The randomization through cloning/mutation processes is fundamental to alter the deterministic nature of standard learning processes introducing robustness to unknown situations. This characteristic is crucial when contextualized in highly heterogeneous scenarios. Third, the process of storage of a small amount of memory-B-cells allows to extract and to maintain the template of what has been learnt ensuring the recognition in case of future encounters. This aspect is fundamental to ensure reproducibility and repeatability. On this basis, standard (AINET) [13] and Adaptive Artificial Immune Networks (A^2INET) [14] have been developed and applied in many diversified scenarios such as, for example, drift mitigation of chemical sensors used in odors classifications [14,15], clinical diagnostic [16], lung cancer detection [17], and breast lesions diagnosis by fine needle aspiration (FNA) biopsy [18]. In standard AINET, the role of the antigen is played by the data (of training or testing) and the antibodies set are the knowledge constructed around the antigens optimizing the criteria of highest affinity and minimum redundancy. The cloning/mutation process results in slight variations of the antibodies ensuring the identification of testing data even subtly different from those used in train. Interestingly, A^2INET has shown a strong capability to detect anomalous data. This feature prompted us to investigate the capability of A^2INET to detect anomalies in data describing mammographic features. Indeed, anomalies in diagnostic image appearance describe any kind of dysfunction or disorder occurring in the normal functioning of the organism. Imaging techniques are simple and noninvasive and can be applied to screen large strata of population. Databases of diagnostic related images are widely available to test classification algorithms, among them mammographic images are particularly interesting because of the importance of an early detection of breast cancer and the existing programs of wide population screening.

Cancer is the most disruptive alteration of living beings affecting a wide variety of body compartments [19]. Breast cancer, in particular, is the leading type of cancer in women, accounting for 25% of all cases [20]. In 2012 it resulted in 1.68 million cases and 522.000 deaths. It is more common in developed countries and about 100 times more frequent in women than in men. Early diagnosis and treatment are the keys for the improvement of quality of life and mortality reduction. The goal of breast screening exams is the early detection of cancer in asymptomatic women where no felt lump or palpable masses are present. In early stages, breast tumors are small and confined in a breast mass. Sometimes, breast cancer is signaled by an asymmetric or uneven distribution of the left and right breast tissues [21–23]. On the other hand, sensitivity and specificity of screening mammography are still relatively low. Misinterpretation may be due to the large variability of breast appearance and abnormalities, or to the overlapping of tissue on the projection images.

To facilitate the correct interpretation of mammography, in the last decade, several computerized systems have been designed, and made available for clinical practice as second readers [24]. In this context, predicting cancer risk for a woman has gained increasing interest under the new paradigm to develop personalized protocols and follow-up for women with different risks of developing cancer [25]. Previous studies have shown that the difference of breast size [26] and average density between the left and right breasts [27], later called Bilateral Breast Asymmetry (BBA), might be a risk indicator of developing breast cancer in a near-term. The BBA is interpreted as a phenotype variation of a genotype abnormality (that breaks the paired morphological traits) that may lead to cancer development. Unfortunately, clinical studies have reported that asymmetry accounts for 3–9% of breast cancer cases incorrectly reported by radiologists as showing no evidence of tumor [28], proving that asymmetric distribution of fibroglandular density is a common source of erroneous interpretation [29]. One of the main pitfall of the existing computerized systems for the detection of BBA is likely due to the fact that they are trained and tested on the same dataset by using cross-validation procedures [30–34]. This leads to a poor capability of recognizing abnormalities in the highly heterogeneous and distributed scenario of screening mammography [35]. To improve the accuracy in a realistic scenario it is necessary to deal with aspects such as the scarcity of abnormalities and the need to distinguish physiological differences in the breast tissue from pathological BBA. An optimal classifier should be able to extract the salient information from a given dataset of mammograms and then identify the abnormalities eventually present in a different dataset. In clinical practice, in fact, the testing mammograms may be acquired, for example, with a different equipment and by using different general device settings, thus generating variations in terms of pixel depth, spatial resolution, and image contrast [36].

In this paper, we show that an algorithm based on an ensemble of A^2INET can capture the signature of asymmetry thanks to the extraction of specific similarity descriptors computed locally on paired mammographic regions [37] which were matched using the localization of the nipple, the breast skin line, and the pectoral muscle in each mammogram [38–40]. A key factor in our strategy is that the selected indicators emulate radiologist's perception of bilateral parenchymal distortions in mammograms. Moreover, the adaptive aptitude of A^2INET allows recognizing different kind of asymmetric findings (e.g. focal or global) in images with intrinsically different appearance (projection view, acquisition, and digitalization settings). The use of an ensemble of classifiers increases the robustness of the model and reduces the dependence of the performance from the model initialization, which is typical in stochastic or non-linear algorithms. In such a way, the non-deterministic nature of the A^2INET algorithm is reduced in favor of greater robustness and repeatability of the results. The case studies here considered bring together two publicly available datasets of mammograms, the Digital Dataset for Screening Mammography (DDSM) [41] and the Mammographic Image Analysis Society digital mammogram database (mini-MIAS) [42], and for both the sets, all the asymmetric cases were considered. The combination of similarity descriptors and A^2INET ensemble results in good classification rates with very low number of false negatives. Remarkably, the performance is preserved when the model is trained with one dataset and tested with another dataset. These results are very encouraging for the perspective of generalizing the results obtained for the automatic detection of asymmetry using a network of different instrument databases. The paper is organized as follows. In Section 2, we illustrate the innovative A^2INET system and in