

# A Deep Learning approach to Case Based Reasoning to the Evaluation and Diagnosis of Cervical Carcinoma

José Neves<sup>1</sup> [00351-934201337], Henrique Vicente<sup>1,2</sup>, Filipa Ferraz<sup>1</sup>, Ana Catarina Leite<sup>3</sup>, Ana Rita Rodrigues<sup>3</sup>, Manuela Cruz<sup>3</sup>, Joana Machado<sup>4</sup>, João Neves<sup>5</sup>

<sup>1</sup> Centro Algoritmi, Universidade do Minho, Braga, Portugal

<sup>2</sup> Departamento de Química, Universidade de Évora, Évora, Portugal

<sup>3</sup> Departamento de Informática, Universidade do Minho, Braga, Portugal

<sup>4</sup> Farmácia de Lameiras, Braga, Portugal,

<sup>5</sup> Mediclinic Arabian Ranches, PO Box 282602, Dubai, United Arab Emirates  
jneves@di.uminho.pt, hvicente@uevora.pt, filipatferraz@gmail.com,  
{a82768, a75951, a79392}@alunos.uminho.pt,  
{joana.mmachado, joaocpneves}@gmail.com

**Abstract.** *Deep Learning (DL)* is a new area of *Machine Learning* research introduced with the objective of moving *Machine Learning* closer to one of its original goals, i.e., *Artificial Intelligence (AI)*. *DL* breaks down tasks in ways that makes all kinds of machine assists seem possible, even likely. Better preventive healthcare, even better recommendations, are all here today or on the horizon. However, keeping up the pace of progress will require confronting currently *AI*'s serious limitations. The last but not the least, *Cervical Carcinoma* is actually a critical public health problem. Although patients have a longer survival rate due to early diagnosis and more effective treatment, this disease is still the leading cause of cancer death among women. Therefore, the main objective of this article is to present a *DL* approach to Case Based Reasoning to the evaluation and diagnosis of *Cervical Carcinoma* using *Magnetic Resonance Imaging*. It will be grounded on a dynamic virtual world of complex and interactive entities that compete against one another in which its aptitude is judged by a single criterion, the *Quality of Information* they carry and the system's *Degree of Confidence* metric on such a measure, under a *fixed symbolic structure*.

**Keywords:** Artificial Intelligence; Deep Learning; Machine Learning; Cervical Carcinoma; Magnetic Resonance Imaging; Logic Programming; Knowledge Representation and Reasoning; Case Based Reasoning.

## 1 Introduction

Improving *Patient Care* with *Artificial Intelligence (AI)* is transforming the world of *Medicine*. *AI* can help doctors make faster, more accurate diagnosis or predict the risk of a disease in time to prevent it. Indeed, although *AI* has been around for decades, new advances have ignited a boom in the *AI* field, namely due to the advent of *Deep Learning (DL)* [1, 2]. This *AI* technique has been powering self-driving cars, image

recognition, and even life-saving advances in *Medicine*, just to name a few. Undeniably, *DL* helps researchers analyze medical data to treat diseases. It is advancing the future of *Personalized Medicine*. But the peculiar thing with the present approach to *DL* is just how old its ideas are. Especially, we wonder if we are at the beginning of a revolution; if a real intelligence breaks when slightly someone changes the problem or if we are endorsing the case where *AI* is riding a one-trick pony? [3].

On the other hand cervical cancer still is portrayed as the second most common one occurring in females [4], with 70% of cases up in developed countries [5] (e.g., in 2013, 11 955 women in the United States were diagnosed with cervical cancer, of which 4 217 died [6]). This type of cancer arises in the cervix, resulting in an abnormal growth of cells with the ability to invade or spread to other parts of the body. It comes into two major sub-types, namely *Cervical Squamous Cell Carcinoma (CSCC)* and *Endocervical Adenocarcinoma (EA)* [7].

Last but not the least, *Magnetic Resonance Imaging (MRI)* is used in medical imaging field to visualize in detail internal structures of the body. Compared with other medical imaging techniques such as *Computer Tomography (CT)* and *X-rays*, it provides good contrast between the soft tissues of the body. It is also used to visualize deformed tissues present in bones, teeth and even fossils [8]. In this study *MRI* was used to evaluate the stage of the *CSCC* of a given patient. Features of resonance images, such as *Tumor Volume*, were extracted using a dataset of 54 patient images that stand for previously experienced and studied concrete situations (cases). One's approach to problem solving will be grounded on *Case Based Reasoning (CBR)*, where a new problem is solved by reporting to a similar past case and reusing its found solution [9]. It is under this setting that we intend to answer the questions that have been referred to above.

This paper is subdivided into five sections, being the former one a part where the problem under analysis is fixed, followed by a background's one where issues related to it are open. The third section introduces the time-line of the problem solving process. The next one addresses the way one comes to a solution using *CBR*. Finally, a conclusion is presented and directions for future work are outlined.

## 2 Background

### 2.1 Cervical Squamous Cell Carcinoma

*Cervical Cancer (CC)* is the second most common cancer in women, being only surpassed by *Breast Cancer* [10]. *CC* is located in the lower part of the uterus and, like other cancers, is related to the abnormal growth of cells that have the capacity to invade other parts of the body [7].

*CC* presents two different parts and is lined by two types of cells, namely glandular cells in the area closest to the body of the uterus and termed the endocervical, and squamous ones lining the part closest to the vaginae (exocervical) [7].

Depending on the cells lining the cervical, there are also different types of cancer, i.e., there are squamous cell carcinoma with cancer cells, adenocarcinoma developed from glandular cells and, less frequently, adenosquamous or mixed carcinomas be-

cause they have characteristics of carcinomas of squamous cells and adenocarcinomas [7]. Here, *CSCC* was the only one studied. Some of the major risk factors for the development of *CSCC* are related to numerous factors, ranging from *Human Papilloma-virus (HPV)* infection, smoking, multiple sexual partners, a weakened immune system, *HIV* infections or organ transplantation [11].

Once the carcinoma has been identified, it is evaluated the way it has spread, thus determining the staging, as it is depicted in Table 1.

**Table 1.** Cervical Carcinoma's different stages [12].

Stage	Extent of disease	5-year survival
0	Carcinoma <i>IN situ</i> ( <i>CIN</i> )	~100%
I	Limited to cervix	
Ia1	Microscopic disease: stromal invasion <3mm, lateral spread <7mm	>95%
Ia2	Microscopic disease: stromal invasion <3mm and >5mm, lateral spread <7mm	
Ib1	Macroscopic lesion <4cm in greatest dimension	~90%
Ib2	Macroscopic lesion >4cm in greatest dimension	80-85%
II	Extension to uterus/parametria/vagina	~75-78%
IIa1	Involvement of upper two thirds of vagina <i>without</i> parametrial invasion, <4cm greatest diameter	
IIa2	Involvement of upper two thirds of vagina <i>without</i> parametrial invasion, >4cm greatest diameter	
IIb1	Involvement of upper two thirds of vagina <i>with</i> parametrial invasion	
III	Extension to pelvic side wall and/or lower third of vagina	~47-50%
IIIa	Involvement of lower third of vagina	
IIIb	Extension to pelvic side wall and/or hydronephrosis	
IV	Extension to adjacent organs or beyond true pelvis	~20-30%
IVa	Extension to adjacent organs e.g. bladder, bowel	
IVb	Distant metastases	

## 2.2 Deep Learning versus Knowledge Representation and Reasoning

Many approaches to integrate *Deep Learning (DL)* with *Knowledge Representation* and *Reasoning (KRR)* are based on the fact that one must give up on having a *fixed symbolic structure* to it [1]. However, when working on such a *relaxation* process, in situations where the *KRR* systems are *induced* by learning algorithms (i.e., Rumelhart, D.E., Hinton, G.E., Williams *DL* [1]), the process turns out to be mostly opaque to the programmers [2]. Putting the things in this form, the distinctiveness about *DP* is just how old its ideas are. This stands for the key distinction between the former approaches (in which it is asserted that the work done is symbolic logic in vector spaces, remaining *discrete* the essential features, and nothing is gained), to the one that will be presented here that, although having a symbolic logic in vector spaces, the elements or attributes of the logical functions there described go from discrete to contin-

uous, allowing for the representation and handling of *unknown*, *incomplete*, *forbidden* and even *self-contradictory information* or *knowledge*.

Definitely, there are many approaches to *KRR* using the *Logic Programming (LP)* epitome, namely in the area of *Model Theory* [13, 14] and *Proof Theory* [15, 16]. Here, the *Proof Theoretical* approach in terms of an extension to the *LP* language is followed. An *Extended Logic Program* is, therefore, given by a finite set of clauses, in the form:

$$\{$$

$$\neg p \leftarrow \text{not } p, \text{not exception}_p$$

$$p \leftarrow p_1, \dots, p_n, \text{not } q_1, \dots, \text{not } q_m$$

$$?(p_1, \dots, p_n, \text{not } q_1, \dots, \text{not } q_m) \quad (n, m \geq 0)$$

$$\text{exception}_{p_1}, \dots, \text{exception}_{p_j} \quad (0 \leq j \leq k), \text{ being } k \text{ an integer number}$$

$$\} :: \text{scoring}_{value}$$

**Program 1.** The Archetype of a Generic Extended Logic Program

where the first clause stand for predicate's closure, “,” denotes “*logical and*”, while “?” is a domain atom denoting falsity, the  $p_i$ ,  $q_j$ , and  $p$  are classical ground literals, i.e., either positive atoms or atoms preceded by the classical negation sign  $\neg$  [16]. Indeed,  $\neg$  stands for a strong declaration that speaks for itself, and *not* denotes *negation-by-failure*, or in other words, a flop in proving a given statement, once it was not declared explicitly. Under this formalism, every program is associated with a set of *abducibles* [13, 14], given here in the form of exceptions to the extensions of the predicates that make the program, i.e., clauses of the form:

$$\text{exception}_{p_1}, \dots, \text{exception}_{p_j} \quad (0 \leq j \leq k), \text{ being } k \text{ an integer number}$$

that stand for data, information or knowledge that cannot be ruled out. On the other hand, clauses of the type:

$$?(p_1, \dots, p_n, \text{not } q_1, \dots, \text{not } q_m) \quad (n, m \geq 0)$$

also named *invariants*, allows one to set the context under which the universe of discourse has to be understood. The term  $\text{scoring}_{value}$  stands for the relative weight of the extension of a specific predicate with respect to the extensions of peers ones that make the inclusive or global program.

In order to evaluate the data or knowledge's assets that may be associated with a logical program, an assessment of it is given in terms of the *Quality-of-Information*

(*QoI*) and *Degree-of-Confidence* (*DoC*) metrics, that range between 0 and 1 and lead to predicates or logical functions in the form [17-19]:

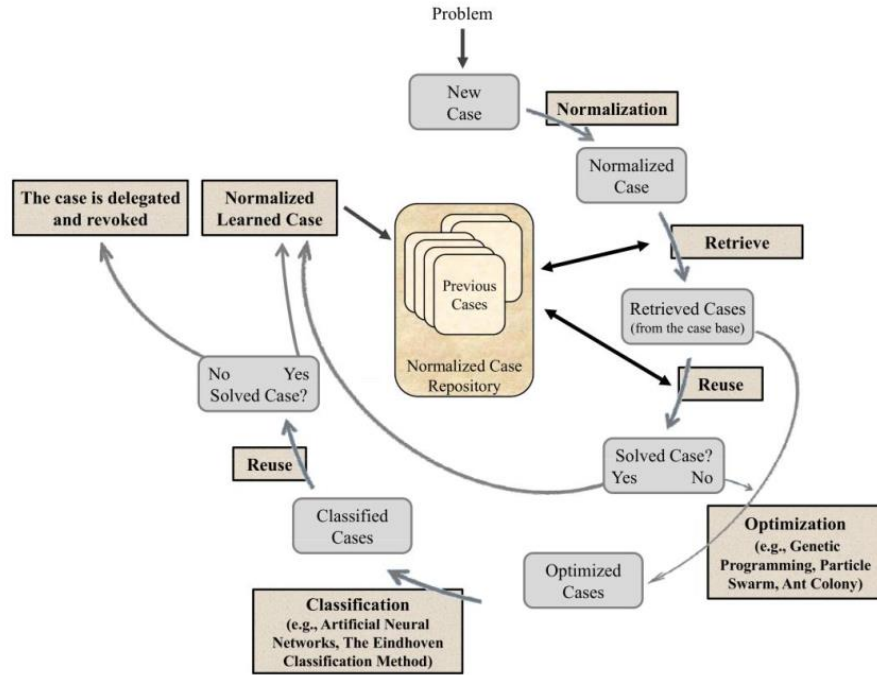
$$predicate_{new} = \bigcup_{1 \leq j \leq m} new_j \left( ([A_{x_1}, B_{x_1}](QoI_{x_1}, DoC_{x_1})), \dots \right. \\ \left. \dots, ([A_{x_n}, B_{x_n}](QoI_{x_n}, DoC_{x_n})) \right) :: QoI_j :: DoC_j$$

which answer our question stated above, i.e., *AI* is no more riding a one-trick pony.

### 2.3 Case Based Reasoning

In order to represent the *CBR* process, we will focus on an extended version of the *CBR* cycle (Fig. 1) [20, 21], an extension of the conventional one [9, 22].

The former step comprises an initial description of the problem or new case, where its characteristics and attributes are identified. Then, in the normalization phase, the new case feature vector attributes' values are committed to values in the range [0, 1]. Then, the retrieved cases from the case base are serial in order to highlight the ones that best overlap with those of the new case [20, 21].



**Fig. 1.** The updated view of the *CBR* cycle proposed by Neves *et al.* [20, 21].

## 2.4 Pre-processing and Segmentation

Pre-processing aims to improve the visual appearance of images and improve the manipulation of datasets. To conduct this process the *SimpliFilters* module of *3D Slicer* was used, which provides a simple interface for hundreds of basic and advanced *ITK* filters. In this module, a large number of filters were found, and in this study *AddImageFilter* was used, which allowed the addition of pixels of two images, being the most frequently used the *BinaryCoutourImageFilter*, which marks the pixels at the edge of objects [23]. The algorithms provided included binary morphology, grayscale morphology, weighting, threshold, manipulation of image intensity, growing region, *Fast Fourier Transform*, just to name a few.

Image segmentation is the process of dividing an image into sub regions using discontinuity and similarity properties, such as the gray level and the precise definition of its spatial extent, to which there exists three targeting techniques, i.e., the region-based, contour and texture-based segmentation [24-26].

## 2.5 Feature Extraction

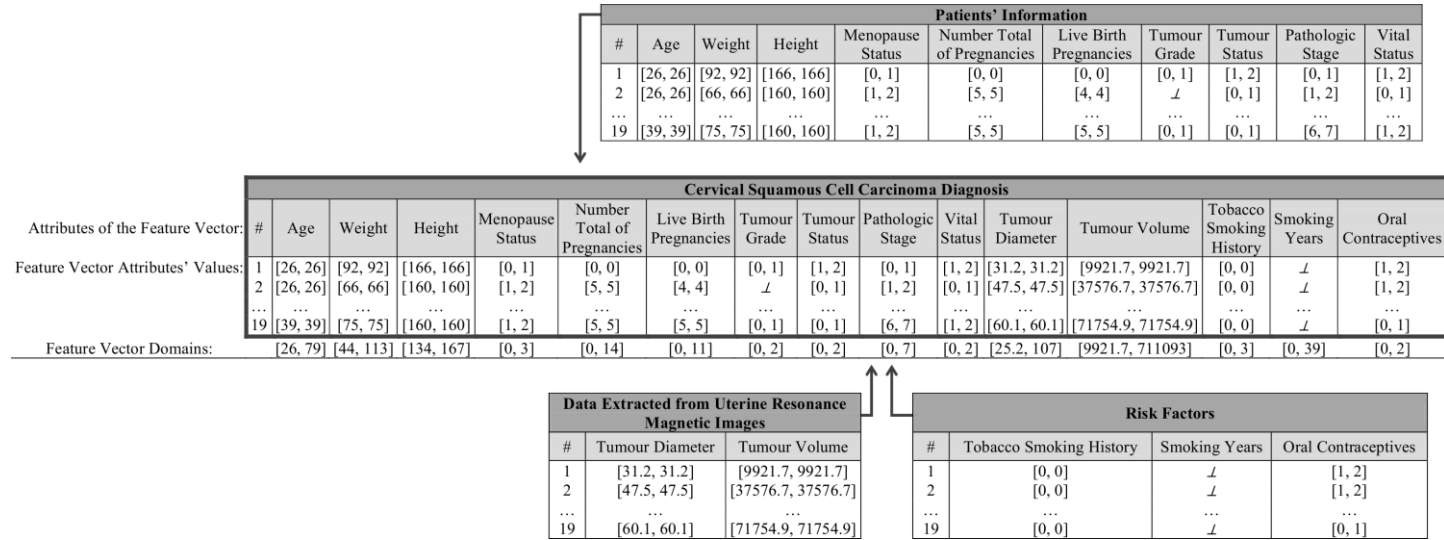
At this stage and in terms of the feature set the following issues were considered:

- *Tumor Volume*; and
- *Tumor Diameter*.

Once having accomplished the segmentation step, the *3D Slicer Ruler* was used to measure the diameter as well as the *LabelStatistics* module that allowed the calculation of the tumor's volume. These extracted characteristics will produce a feature vector – VC – that will represent the dataset. Other characteristics, such as *Age*, *Weight*, or *Risk Factors* were extracted in the form of a text file from the image repository *The Cancer Imaging Archive (TCIA)* [27].

## 3 Case Study

After choosing the sample of patients to be studied and describing the procedures that served as a basis for the pre-processing and extraction of the *MRI* features, it were constructed tables that present the values or its ranges to each attribute studied for the diagnosis of Cervical Squamous Cell Carcinoma, namely the patient's data, the *Risk Factors* associated with the diagnosis, and the data extracted from the *MRI* images (Fig. 2). Thus, it is now possible to set the objective function to the problem under analysis in terms of the extension of predicate *diag<sub>CSCC</sub>* [19]. In the emergency of a new case, for instance the one with feature vector *Age* = 52; *Weight* = 66; *Height* = 156; *Menopause Status* = [0, 1], *Number Total Pregnancies* = 2, *Live Birth Pregnancies* = 1, *Tumor Grade* = 1, *Tumor Status* = [1,2], *Pathologic Stage* = [6, 7], *Vital Status* = [0, 1]; *Tumor Diameter* = 0.4670; *Tumor Volume* = 0.0650; *Tobacco Smoking History* = 1; *Smoking Years* = 1; *Oral Contraceptives* = [0, 1], one may have:



**Fig. 2.** Knowledge Base for the Diagnosis of Cervical Squamous Cell Carcinoma.

$$diag_{CSCC_{NEW CASE}} \left( ((0.49, 0.49)(1, 1)), \dots, ((0, 0.5)(1, 0.87)) \right) :: 1 :: 0.84$$

The new case is now compared with each case retrieved from the case base, and using as similarity function the mean of the module of the arithmetic difference between the arguments of each selected case and the arguments of the new case, one may get:

$$\begin{aligned} & retrieved_{case_1} \left( ((0.58, 0.58)(1, 1)), \dots, ((0, 0.5)(1, 0.87)) \right) :: 1 :: 0.96 \\ & \vdots \\ & retrieved_{case_n} \left( ((0.47, 0.47)(1, 1)), \dots, ((0, 1)(1, 0)) \right) :: 1 :: 0.84 \end{aligned}$$


---

*normalized cases that make the retrieved cluster*

Assuming that each attribute has equal weight, the dissimilarity, the *DoC*'s similarity between the new case and the former retrieved one (*retrieved case<sub>1</sub>*) is evaluated in the form:

$$dis_{NEW CASE \rightarrow 1}^{DoC} = \frac{|1 - 1| + \dots + |0.87 - 1|}{15} = 0.022$$

leading to:

$$sim_{NEW CASE \rightarrow 1}^{DoC} = 1 - 0.022 = 0.978$$

The same process is applied to the *QoI*, leading to the overall similarity:

$$sim_{NEW CASE \rightarrow 1}^{QoI, DoC} = 1 \times 0.978 = 0.978$$

These procedures may be extended to all retrieved cases leading to the most similar ones and the potential solutions to the problem returned to the physicians [27].

The model's performance, namely in terms of sensitivity (81.4%) and specificity (80.3%) denotes that the proposed model exhibits an acceptable performance in the diagnosis of cervical squamous cell carcinoma.

## 4 Conclusions

The purpose of this paper was to present a framework to construct a dynamic virtual world of complex and interactive entities that map real cases of *Cervical Squamous Cell Carcinoma* in order to develop a decision support system to help predict the different stages of this disease. Based on a new *DL* approach to *KRR*, it also proved that when you boil it down, *AI* may be close to *DL*, but *DL* is not *Backprop*. Indeed, it was described the extraction of features that allowed the classification of stages of the *Cervical Squamous Cell Carcinoma* through a system of segmentation and classification based on a simple and fixed symbolic structure, that sets a new understanding to *DL*. This leads us to the construction of an inductive theory using case studies, speci-



ifying the research questions for closure, in the field of Carcinoma evaluation. The basic idea of the dynamic mechanism of evolutionary heuristics follows the basic rule of the *CBR* approach to evolution, where the learning procedure is based on the evolution of knowledge and is constructed based on a process of *QoI* and DoC's metrics quantification that results from evolving *Logical Programs* or *Theories*.

**Acknowledgments.** This work has been supported by COMPETE: POCI-01-0145-FEDER-007043 and FCT – Fundação para a Ciência e Tecnologia within the Project Scope: UID/CEC/00319/2013.

## References

1. Rumelhart, D.E., Hinton, G.E., Williams, R.J.: Learning Internal Representation by Error Propagation. In: Rumelhart, D.E., McClelland, J.L. (eds.) *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, vol. 1. pp. 318–362, MIT Press, Cambridge (1986)
2. Rumelhart, D.E., Hinton, G.E., Williams, R.J.: Learning representations by back-propagating errors. *Nature* 323, 533–536 (1986)
3. Sommers, J.: Is AI Riding a One-Trick Pony?. <https://www.technologyreview.com/s/608911/is-ai-riding-a-one-trick-pony/>
4. Sousa, A.: *Cervical Cancer: Trends and Studies*. Fernando Pessoa University Edition, Oporto (2011)
5. Deng, S., Zhu, L., Huang, D.: Predicting Hub Genes Associated with Cervical Cancer through Gene Co-Expression Networks. *IEEE/ACM Transactions on Computational Biology and Bioinformatics* 13, 27–35 (2016)
6. U.S. Cancer Statistics Working Group: United States Cancer Statistics: 1999–2014 Incidence and Mortality Web-based Report. <http://www.cdc.gov/uscs>
7. American Cancer Society: What is cervical cancer?. <https://www.cancer.org/cancer/cervical-cancer/about/what-is-cervical-cancer.html>
8. Remya, V., Lekshmi-Priya V.L.: Simultaneous Segmentation and tumor detection in MRI cervical Cancer Radiation Therapy with hierarchical adaptive local affine registration. *Proceedings of the 2014 International Conference on Computer Communication and Informatics*, 6 pp. IEEE Edition (2014)
9. Aamodt, A., Plaza, E.: Case-based reasoning: Foundational issues, methodological variations, and system approaches. *AI Communications* 7, 39–59 (1994)
10. Mithlesh, A., Namita, M., Girdhari, S.: Cervical cancer detection using segmentation on Pap smear images. *Proceedings of the 1<sup>st</sup> International Conference on Informatics and Analytics*, Article 29, Association for Computing Machinery, New York (2016)
11. Vaccarella, S., Lortet-Tieulent, J., Plummer, M., Franceschi, S., Bray F.: Worldwide trends in cervical cancer incidence: Impact of screening against changes in disease risk factors. *European Journal of Cancer* 49, 3262–3273 (2013)
12. Pecorelli, S.: Corrigendum to “Revised FIGO staging for carcinoma of the vulva, cervix, and endometrium”. *International Journal of Gynecology and Obstetrics* 108, 176 (2010)
13. Kakas, A., Kowalski, R., Toni, F.: The role of abduction in logic programming. In: Gabbay, D., Hogger, C., Robinson, I. (eds.) *Handbook of Logic in Artificial Intelligence and Logic Programming*, vol. 5, pp. 235–324. Oxford University Press, Oxford (1998)

14. Pereira, L., Anh, H.: Evolution prospection. In: Nakamatsu, K. (ed.) *Studies in Computational Intelligence*, vol. 199, pp. 51–64. Springer, Berlin (2009)
15. Neves, J., Machado, J., Analide, C., Abelha, A., Brito, L.: The halt condition in genetic programming. In: Neves, J., Santos, M.F., Machado, J. (eds.) *Progress in Artificial Intelligence*. LNAI, vol. 4874, pp. 160–169. Springer, Berlin (2007)
16. Neves, J.: A logic interpreter to handle time and negation in logic databases. In: Muller, R., Pottmyer, J. (eds.) *Proceedings of the 1984 annual conference of the ACM on the 5th Generation Challenge*, pp. 50–54. Association for Computing Machinery, New York (1984)
17. Machado J., Abelha A., Novais P., Neves J., Neves J.: Quality of service in healthcare units. In Bertelle, C., Ayesh, A. (eds.) *Proceedings of the ESM 2008*, pp. 291–298. Eurosis – ETI Publication, Ghent (2008)
18. Fernandes, A., Vicente, H., Figueiredo, M., Neves, M., Neves, J.: An Adaptive and Evolutionary Model to assess the Organizational Efficiency in Training Corporations. In: Dang, T.K., Wagner, R., Küng, J., Thoai, N., Takizawa M., Neuhold, E. (eds.) *Future Data and Security Engineering, Lecture Notes on Computer Science*, vol. 10018, pp. 415–428, Springer International Publishing, Cham (2016)
19. Fernandes, F., Vicente, H., Abelha, A., Machado, J., Novais, P., Neves J.: Artificial Neural Networks in Diabetes Control. In *Proceedings of the 2015 Science and Information Conference (SAI 2015)*, pp. 362–370, IEEE Edition (2015)
20. Quintas, A., Vicente, H., Novais, P., Abelha, A., Santos, M.F., Machado, J., Neves, J.: A Case Based approach to assess Waiting Time Prediction at an Intensive Care Unity. In: Arezes, P. (ed.) *Advances in Safety Management and Human Factors. Advances in Intelligent Systems and Computing*, vol. 491, pp. 29–39, Springer International Publishing, Cham (2016)
21. Silva, A., Vicente, H., Abelha, A., Santos, M.F., Machado, J., Neves, J., Neves, J.: Length of Stay in Intensive Care Units – A Case Base Evaluation. In: Fujita, H., Papadopoulos, G. A. (eds.) *New Trends in Software Methodologies, Tools and Techniques, Frontiers in Artificial Intelligence and Applications*, vol. 286, pp. 191–202, IOS Press, Amsterdam (2016)
22. Richter, M.M., Weber, R.O.: *Case-Based Reasoning: A Textbook*. Springer, Berlin (2013)
23. 3D Slicer. <https://www.slicer.org/wiki/Documentation/Nightly/Modules/SimpleFilters>
24. Perdigão, N., Tavares, J.M., Martins, J.A., Pires, E.B., Jorge, R. M.: Sobre a Geração de Malhas Tridimensionais para fins computacionais a partir de imagens médicas. In Aparicio, J.L., Ferran, A.R., Martins, J.A., Gallego, R., Sá, J.C. (eds.) *Proceedings of the Congreso de Métodos Numéricos en Ingeniería*, 16 pp, SEMNI, Barcelona (2005)
25. Hasan, D.I., Enaba, M.M., El-Rahman, H.M.A., El-Shazely, S.: Apparent diffusion coefficient value in evaluating types, stages and histologic grading of cancer cervix. *Egyptian Journal of Radiology and Nuclear Medicine* 46, 781–789 (2015)
26. Camisão, C., Brenna, C., Lombardelli, S., Djahjah , K., Zeferino, L.: Magnetic resonance imaging in the staging of cervical cancer. *Radiologia Brasileira* 40, 207–215 (2007)
27. The Cancer Imaging Archive (TCIA). <http://www.cancerimagingarchive.net/>
28. Faria, R., Alves, V., Ferraz, F., Neves, J., Vicente, H., Neves, J.: A Case Base Approach to Cardiovascular Diseases using Chest X-Ray Image Analysis. In: van den Herik, J., Rocha, A.P., Filipe, J. (eds.) *Proceedings of the 9th International Conference on Agents and Artificial Intelligence (ICAART 2017)*, Vol. II, pp. 266–274, Scitepress – Science and Technology Publications, Lisbon, (2017)