SURVEY OF MACHINE LEARNING APPLICATIONS IN MEDICAL IMAGING. ALGORITHMS AND TECHNOLOGIES

Eng. PhD. Marius Eremia Vlaicu POPA, Complexul Energetic Oltenia, ROMANIA, m.vlaicu@yahoo.com
Professor PhD. eng. Mihai CRUCERU, Constantin Brancusi" University of Tg-Jiu, ROMANIA, cruceru.mihai@gmail.com
Assoc.Prof. Eng. Ph.D., Bogdan DIACONU, Constantin Brancusi" University of Tg-Jiu, ROMANIA, bdiaconu2004@gmail.com

Abstract: *A literature survey was conducted attempting to report usage of Machine Learning (ML) algorithms in processing and interpretation of medical images (X-ray, CT scan, MRI, DEXA, etc.) for the purpose of diagnosing orthopedic disorders. ML algorithms and technologies developed exponentially over the last decades triggered by advances in processing capacity, open-source ML platforms, frameworks and libraries (Anaconda, Scikit, TensorFlow, OpenNN, Torch and so on).*
Keywords: Artificial Intelligence; Machine Learning; Medical Imaging; Medical Diagnosis.

1. Introduction

Medical imaging has become a standard in clinical analysis of orthopedic disorders, either trauma, developmental or induced by factor such as infectious processes, neuromuscular, nutritional, or neoplastic processes. A large number of imaging techniques is required in orthopedic diagnosis, depending on the particular tissue that has to be visualized and the specific type of tissue abnormality suspicion. Some case-specific conditions could impose restrictions on the use of one technique or another: patients with cardiac stimulation devices prevents the usage of MRI; various allergies prevent imaging techniques that require contrast agents, etc.

Picture Archiving and Communication Systems (PACS) have been used since the 1990s for radiologic images storage, management and processing [224]. A complex system aiming at storing, transmitting, retrieving, printing, processing and displaying medical imaging information - Digital Imaging and Communications in Medicine was introduced in 1993 0. The necessity to archive/retrieve and search for relevant information in radiology reports resulted in definition of lexicons such as Metathesaurus 0, RadLex 0 and Medical Subjects Headings 0. Metathesaurus includes more than five million concept names and a million biomedical terms from more than one hundred controlled vocabulary systems 0. Such databases require automated agents to add/retrieve/search/manage information, which is where ML models can prove their utility. DL is a ML subset that demonstrated promising results in generating radiology reports from images 0, 0. DL applications in healthcare have been reviewed in 0.

ML has been increasingly used in medical application, especially in diagnosis, with the role of a decision support tool designed to improve the performance of the healthcare provider. The first reports of AI applications in medicine dates back in the 1980s 0. Picciali et al 0 reviewed comprehensively the evolution of DL application in medicine reporting: (1) medical specialties such as oncology, cardiovascular medicine, orthopedics, neurology, pulmonology, etc., (2) bio signals, such as electrocardiography, electroencephalography, phonocardiography, photoplethysmography, electromyography, magnetic resonance spectroscopy and monitoring signaling molecules and (3) DL models such as Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN), Autoencoders, Generative Adversarial Networks (GAN), Deep Belief Networks (DBN).

2. Literature review: Results.

The machine learning technologies identified in the selected papers were classified in the following groups:

1. Neural Networks
2. Support Vector Machine
3. Decision Tree/Random Forest/Gradient Boosting
4. Linear Regression/Multiple Least Squares Linear Regression
5. K Nearest Neighbor
6. Naïve Bayes Classifier

3. Reports of typical ML technologies employed in orthopedic imaging

The main ML technologies that were identified in the literature survey presented in Section 2 are discussed and their applications in orthopedic imaging are be presented in the following sections.

3.1. Artificial Neural Networks.

Neural Networks are the dominant ML technique in orthopedic imaging. ANNs are a large family of ML models developed as early as 1940. The fundamental processing unit of an ANN is the artificial neuron. The artificial neuron represents in fact a simple mathematical algorithm. The neuron has n inputs $x_1, x_2, \dots, x_n$ and one output y_k. The algorithm performs a summation on the inputs (with a set of weights applied) and the result is passed as an argument to an activation function, which produces the output.

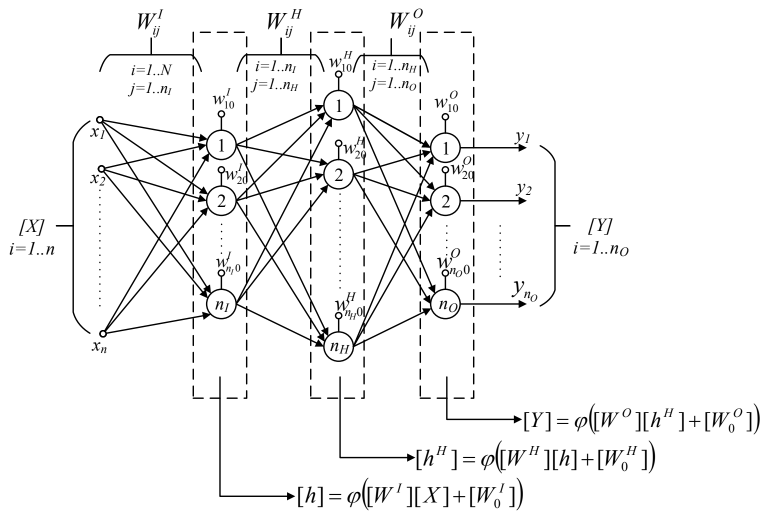

Figure 1. The basic principle of an ANN

This can be written as:

$$y_k = \varphi(w_0 + \sum_{i=1}^{n} w_{ki} x_i) \quad (1)$$

An artificial neural network is obtained by combining more individual units organized in layers, as shown in Figure 1. The two external layers which connect the neural network to the outer world are the input layer (denoted I) and the output layer (denoted O). All layers in between are called hidden layers. Although every individual unit performs a very simplistic operation, the interconnected nature of the system allows very complex calculations and implementation of complex function, which would be impossible to model by means of conventional mathematical means.

A type of ANN frequently encountered in orthopedic imaging for diagnosis purposes is the CNN, represented schematically in Figure 2.

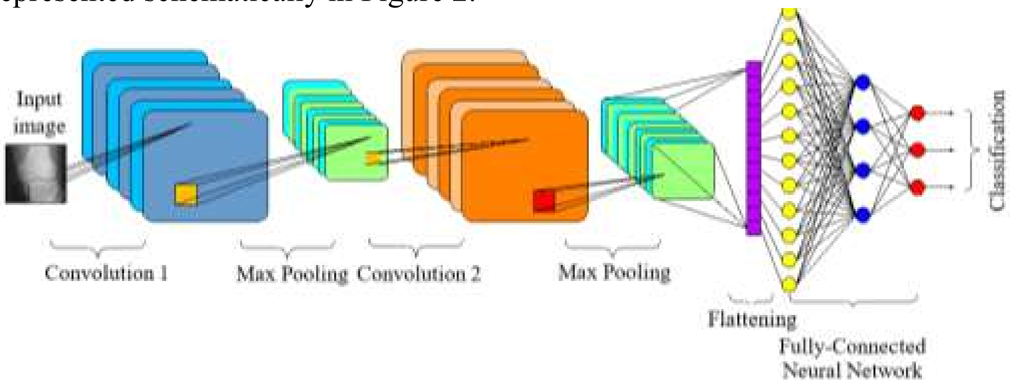

Figure 2. The- CNN architecture

A CNN (or deep CNN) has several features distinguishing it from the standard ANN. The most important, CNNs are deep, in the sense their typical number of layers is 10-30. Another special feature of CNNs is the fact that neurons share weights, which allows the network to perform convolutions of the input image with the filter (defined by the weights). Another distinguishing feature of CNNs is that between some layers, they perform pooling which makes the network invariant to small shifts of the input images. These features make CNNs suitable for image processing.

3.2. Support Vector Machine

SVM are a class of ML technologies that can be used for both regression and classification problems. SVM is a supervised learning type algorithm. It is a discriminative classifier, which identifies a hyperplane that categorizes the points provided as input set. A hyperplane in an n-dimensional space is defined as a *n-1* dimensional space, such as in the two-dimension space the hyperplane is actually a line.

These types of classifiers apply in cases where the training pattern are linearity separable. Consider a hyperplane given by the equation:

[redacted] (2)

where $[\boldsymbol{X}] = [x_1, x_2, \ldots, x_n]$ the hyperplane dimensions and $[\boldsymbol{W}] = [w_1, w_2, \ldots, w_n]$ the hyperplane parameters). The hyperplane separates the space into two halves, positive half space and negative half space. Positive half space satisfies the equality:

[redacted] (3)

and negative half space:

$$[W]\cdot[X] < C \tag{4}$$

SVM is a linear discriminant (we will not discuss the case of points that are not linearly classifiable) that is able to separate two classes of points, as shown in Figure 3.

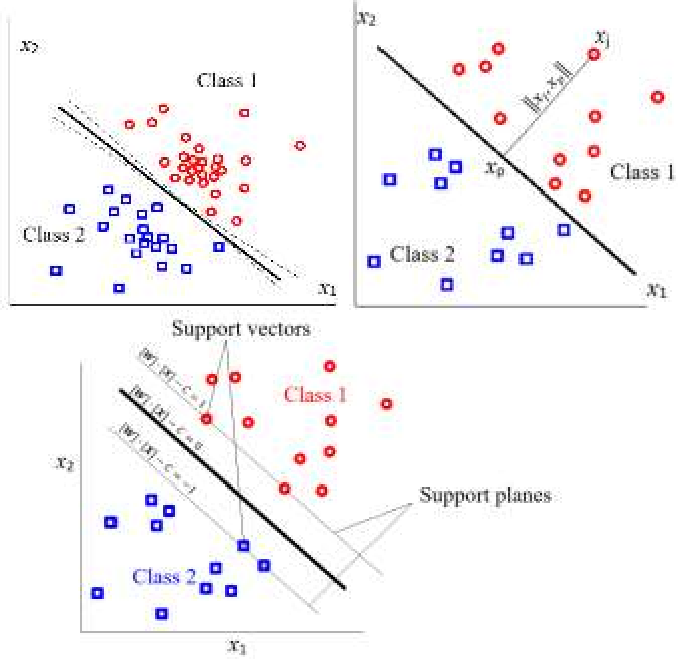

Figure 3. A linear discriminant separating two classes of points

It is clear though that an infinite number of hyperplanes exist that can completely classify the two classes of points (dashed lines in Figure 3). SVM is the decision boundary that maximizes the margin from both patterns and thus creates the best hyperplane that classify the data. To find the minimum distance between the point x_p must be identified in such way that the vector $\overrightarrow{x_j x_p}$ is perpendicular to the decision boundary. This becomes an optimization problem with the following constraints:

$$\left\|x_j, x_p\right\| = min \tag{5}$$

$$[W]\cdot[X_p] = C \tag{6}$$

Condition (5) means that the distance from all points to the decision boundary hyperplane must be minimum. Condition (6) is in fact the hyperplane equation (1) since points $\boldsymbol{X_p}$ are located in the decision boundary hyperplane.

For a two-class problem classified by the hyperplane $[\boldsymbol{W}]\cdot[\boldsymbol{X}] - C = 0$ it can be conventionally considered that the positive class is represented by the hyperplane $[\boldsymbol{W}]\cdot[\boldsymbol{X}] - C = 1$ and the negative class is represented by the hyperplane $[\boldsymbol{W}]\cdot[\boldsymbol{X}] - C = -1$. These hyperplanes are called support planes and points falling on the support planes are called support vectors.

The solution procedure for the optimization problem formulated in equations (5) and (6) is beyond the scope of this paper. The decision boundary hyperplane equation is:

$$[\boldsymbol{X_p}] = [\boldsymbol{X}] - \left(([\boldsymbol{W}][\boldsymbol{X}] - C)\frac{[\boldsymbol{W}]}{\|\boldsymbol{W}\|^2}\right) \tag{7}$$

3.3. Decision Tree and Random Forest

Decision Tree is a supervised learning algorithm used in classification problems. The decision tree consists of nodes connected through branches forming a rooted tree. The decision tree model is a directional one with a node called “root” having no incoming branch. All other nodes have exactly one incoming branch (edge). Any node with outgoing edges is an internal or test node. All other nodes are called leaves or terminal or decision nodes. In a decision tree, each internal node partitions recursively the instance space into two or more sub-spaces according to a certain discrete function of the input attributes values. A major problem of the DT algorithm is that trees are unstable – that is, small changes to the input data can have significant effects on the structure of the tree, due to the hierarchical nature of the tree-growing process, causing errors at the top, which affect dramatically the rest of the tree. One method to reduce the variance of an estimate is to average together several estimates. Training n trees on different subsets of the data, chosen randomly then compute the arithmetic mean is a technique known as Random Forest (RF).

3.4. k-Nearest Neighbor

k-NN is a ML supervised learning algorithm used to classify data by assigning them to the class of the most similar observations. k-NN employs the Euclidian distance to assign a label to any data unit, based on which classification is carried out. Another important concept of the k-NN algorithm is the k value, which determines how many neighbors will grouped in one class. A short but comprehensive introduction to k-NN is given in 0. k-NN has been used in medical image processing for diagnosing orthopedic disorders in 0 (determining location and type of meniscal tears from magnetic resonance images), 0 and 0 (automatic segmentation of knee MRIs to diagnose osteoarthritis).

3.5. Naïve Bayes

Naïve Bayes (NB) is a classifier based on Bayes’ theorem assuming independence between variables and equal contribution to the output. Bayes’ theorem determines the probability of an event occurring given the probability of another event that has already occurred:

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)} \quad (9)$$

where $P(A), P(B)$ are the independent probabilities of events A and B respectively, $P(A|B)$ is the probability of event A, given the event B has occurred, $P(B|A)$ is the probability of event *B*, given the event *A* has occurred.

NB has several advantages over other ML algorithms: it is relatively simple and easy to implement, it dosed not need a large amount of training data, it can handle both continuous and numeric data and it is insensitive to irrelevant features.

4. Conclusions

A literature survey on ML applications in healthcare was performed restricting the studies considered to those related to imaging techniques for diagnosing orthopedic disorders. Selection of algorithms included in this study considered the basic definition of the term Machine Learning [**Eroare! Fără sursă de referință.**], 0, 0.

Three classification criteria were considered in order to perform a systematization of the studies: (1) ML technologies, (2) Imaging technique and (3) Orthopedic disorder type. The systematization according to the criteria mentioned above allowed identification of (a) the ML technology with highest impact on medical imaging in orthopedy; (b) the orthopedic disorders appropriate for application of ML technologies in diagnosis; (3) the medical imaging techniques that can be enhanced/improved/automated by means of ML technologies.

References

[1]. **Maram Mahmoud A. Monshi, Josiah Poon, Vera Chung**, *Deep learning in generating radiology reports: A survey*, Artificial Intelligence in Medicine 106 (2020) 101878

[2]. **Sahu B, Verma R.** *DICOM search in medical image archive solution e-sushrut chhavi*, 2011 3rd International Conference on Electronics Computer Technology, 6. 2011

[3]. **Schuyler PL, Hole WT, Tuttle MS, Sherertz DD**. *The UMLS Metathesaurus: representing different views of biomedical concepts*. Bull Med Libr Assoc 1993;81(2):217

[4]. **Langlotz CP.** *RadLex: A New Method for Indexing Online Educational Materials*, RadioGraphics, vol.26, No.6

[5]. https://www.nlm.nih.gov/mesh/ accesed on 02.02.2021

[6]. **Li CY, Liang X, Hu Z, Xing EP**. *Hybrid Retrieval-Generation Reinforced Agent for Medical Image* Report Generation arXiv preprint arXiv:180508298 2018.

[7]. **Jing B, Xie P, Xing E**. ***On the Automatic Generation of Medical Imaging*** Reports arXiv preprint arXiv:171108195 2017.

[8]. **Andre Esteva, Alexandre Robicquet, Bharath Ramsundar, Volodymyr Kuleshov, Mark DePristo, Katherine Chou, Claire Cui, Greg Corrado, Sebastian Thrun, Jeff Dean**, *A guide to deep learning in healthcare*, Nature Medicine, vol.25, January 2019

[9]. **I. Goodfellow, Y. Bengio, A. Courville, Deep Learning, MIT press, 2016**
Francesco Piccialli, Vittorio Di Somma, Fabio Giampaolo, Salvatore Cuomo, Giancarlo Fortino, *A survey on deep learning in medicine: Why, how and when? Information Fusion 66* (2021) 111–137

[10]. **Zhongheng Zhang**, *Introduction to machine learning: k-nearest neighbors*, Ann Transl Med 2016;4(11):218

[11]. **Ahmet Saygili, Songül Albayrak**, *An efficient and fast computer-aided method for fully automated diagnosis of meniscal tears from magnetic resonance images,* Artificial Intelligence In Medicine 97 (2019)

[12]. **Erik B. Dam, Martin Lillholm, Joselene Marques, Mads Nielsen**, *Automatic segmentation of high- and low-field knee MRIs using knee image quantification with data from the osteoarthritis initiative*, J Med Imaging (Bellingham). 2015 Apr;2(2)

[13]. **Guillaume Madelin, Frederick Poidevin, Antonios Makrymallis, Ravinder R. Regatte,** *Classification of Sodium MRI Data of Cartilage Using Machine Learning*, Magnetic Resonance in Medicine 74:1435–1448 (2015)

[14]. **C.M. Bishop**, *Pattern Recognition and Machine Learning*. Springer 2006

[15]. **Shai Shalev-Shwartz, Shai Ben-David**, *Understanding Machine Learning: From Theory to Algorithms,* Cambridge University Press 2014.