Ideas for Dealing with Reduced Datasets in Development of CADe Systems for Medical Uses

José Morista Carneiro da Silva
Computing Institute
Universidade Federal Fluminense - UFF
Niteroi, Brazil
josemorista@id.uff.br

Aura Conci
Department of Computer Science
Universidade Federal Fluminense - UFF
Niteroi, Brazil
aconci@ic.uff.br

Abstract—This work presents a real time user friendly system to aid specialized professionals to analyze bone scans exams. In order to achieve this, some original ideas are applied. The first one is related to the use of each pixel of an exam as object of interest for classification. Another original idea is the use of operations that are normally applied in pre-processing as features for machine learning. With both, even using small dataset was possible to obtain enough amounts of entries to be used for training and testing. Initially, the feature vectors are composed by 64 features and one target attribute representing the classification result. The used bone scans set was composed of 42 images from 21 patients. At the end of the learning tasks a dataset of 2,512,386 records is computed. In order to reduce the cardinality of the vector of features, the Principal Component Analysis was employed leading to a new feature set with 25 components per object to be classified as with or without metastasis, the area under the Receiver Operator Characteristic curve achieved with this final set of features was 98%.

Keywords—Medical image processing, Bone scintigraphy, e metastasis.

1. Introduction

Nowadays, healthcare industry is one of the most active and important industrial sector. After the last years' reorganization due to the COVID-19 pandemic, it yearns like never before for digital solutions, artificial intelligence tools and innovation [1]. The global health care continues to rise up to the new challenges to produce and procure the required tools for treatments and diagnosis. At the same time, they continue to address the heightened importance of fast outline of the actions to help build improved health care outcomes for all types of diseases. Considering such produce and procure cancer continues to be a very serious cause of dead, especially after the primary treatment. This is mainly due to cancer spread by the body lymphatic system to other locations. Metastases are the activities associated with the occurrence of new secondary cancers at a distance from a primary site [2].

Bone is a frequent site of metastases and usually indicates a short-term prognosis for patients. Once the cancer has spread to the bones, it can rarely be cured, however, it can often still be treated with the aim of slowing their growth and help improve quality of life. Identifying metastatic bones requires a systematic inspection of scintigraphy exams by trained and experienced specialists. However, in the majority of the occasions, this process is performed only through a pure visual analysis, making impossible carry out a quantitative assessment of the issue extension and then an individualized treatment. The report of an examination made on visual bases makes impractical to compare the progression of the patient with the time, because evaluations based on apparent evaluation may not be faithfully reproducible at a future examination done by same person and even less coincident with the evaluation of another professional. The inability to quantify the findings imposed by purely visual analysis directly impacts the routine of specialists, promoting greater difficulty in monitoring the evolution of the disease and consequently, in conducting a proper treatments. To have a way to better detect and quantify bone metastases can significantly improve the treatments, which can be decisive in fighting against cancer spread and vital for maintenance of patients life in most cases.

The bone scan or scintigraphy is a very low-contrast type of imaging. Despite the high expertise a nuclear medicine professional must have, there are limits to the human eye when faced with these gamma radiation images to perform a quantitative count of the findings in an exam. However, computer vision and machine learning tasks can help on correctly classify areas of metastases and measure them. Once the presence of secondary tumors in bone regions can be quantified, it is possible to efficiently measure the advance rate of the affected area, enabling better monitoring the treatment and even modify it. Consequently, providing this computational aid can be a way to bring great benefits to the medical community and patients suffering the effects of this serious condition. Details of a solution usable on desktop or mobile devices are here presented.

2. Bone Scan Exam

In some cancers, where bone metastasis is common, tests might be done for early detection. Some of the tests to allow this diagnosis are: X-ray; computerized tomography (CT); magnetic resonance (MR); positron emission tomography (PET) and bone scan. Among such exams bone scan or bone scintigraphy is one of the most common [3]. It is usable for all gender and becomes the most suitable medium for a definitive decision [4]. A bone scan is a nuclear medicine exam in which a radioactive fluid that is attracted to cancer affected cells is injected into the patient's vein, after a specific scanner takes images of the patient's body and areas of cancer in the bone appear as highlighted areas [5]. This exam is the method of imaging with the highest sensitivity (95%) regarding bone metastases [6]. Figure 1 shows a screen of our implementation where is possible to see on the left image how is the output of this process. Unlike other

exams such as CT and MRI in it only one image from each direction (front or back) is formed, presenting low resolution and very poor contrast [7]. Commonly a full body dorsal or ventral image has 256×1024 pixels. The areas where the radionuclides are concentrated are called ”hotspots” and may indicate the presence of conditions such as arthritis, malignant (cancerous) bone tumors, metastatic bone cancer, bone infections, bone trauma (invisible on ordinary X-rays) and other bone issues [8]. Bone scan images are usually presented with foreground in lighter grays and background in black, however these shades can be inverted to allow better visibility.

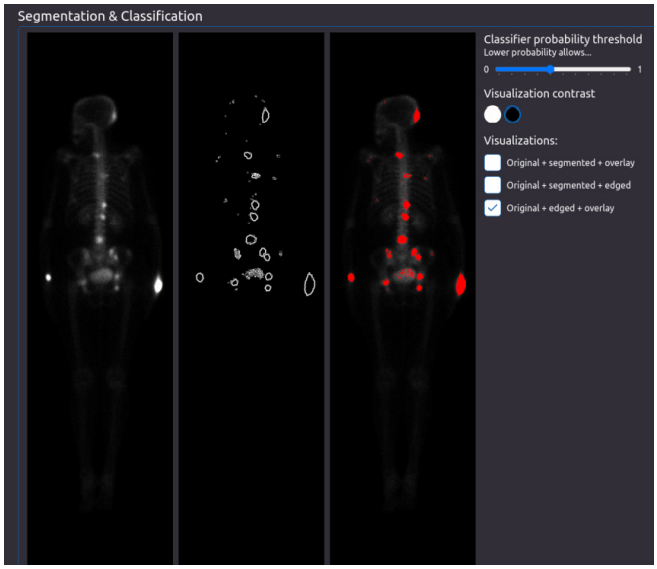

Fig. 1. Screen of the implemented application for metastic identification, with an original image, the border of the hotspots and both overlapped.

3. Previous Works

Analyzing the works carried out on counting of bone metastases, it is important to mention the introduction of Bone Scan Index (BSI) [9]. BSI is a quantitative assessment of bone scan that represents the weight of total tumor as a fraction of the patient's skeletal weight. It has shown clinical utility as a prognostic biomarker to associate therapy and bone density. Despite demonstrating clinical utility, the application and dissemination of the BSI in clinical practice have been limited because of its laborious calculations [10].

The EXINI Diagnostic system used Image Processing (IP) and Artificial Neural Network (ANN) for automating BSI computation [7]. It used a total of 200 patients (100 were used for training and 100 for validation) and a dataset with 400 images [7]. EXINI was further improved using a dataset with 810 patients [8]. In both versions, four IP operations are used: skeleton segmentation, hotspot detection, feature extraction and classification. For segment the skeleton, in first version, simple threshold was use, while the later version this is based on active-shape models (ASM) [8]. In order to detect potential hotspots, the work did simple statistic evaluation of the mean and standard deviation of pixels in a region; when above a threshold, the area is considered hotspot. The area must be equal or greater than 6 pixels in EXINI v1 and it must be bigger than 13 pixels in EXINI v2 to be labeled as potential hotspots [7, 8]. Fourteen features were used in v1, while 45 features are computed in later version. A sensitivity of 90% is achieved for both versions but the specificity improves from of 74% in v1 to 89% using v2.

BONENAVI software also used ANN and was presented in 2 versions. Its first version was trained using 904 patients from a single Japanese institution and after with 1,532 patients from nine institutions [9, 11]. They presents the area under the ROC curve (AUC) that was 0.912 for the first version and 0.934 for the last release [11]. The sensitivity and specificity based on the last version of BONENAVI was 92% [11].

CADBOSS used 130 images from 60 patients, 100 with metastases. Ten folds cross validation was used during the training and test phases [12]. CADBOSS used hotspot segmentation, feature extraction, feature selection and classification. The method used for segmentation was the Level Set Active Contour (LSAC). After segmentation, feature extraction was carried out. The best features was selected using the Principal Component Analysis (PCA). The hotspots are classified by the ANN as presenting or did not presenting metastases. Body images were reduced to 200×700 pixels and a set of sub images were created by dividing each image into 25 equal pieces. Thus, 625 sub images were obtained with 8×28 pixels. Only the ten (10) most important features selected for the PCA were inserted in the ANN. The values of accuracy, sensitivity, and specificity of the software were 92.30%, 94%, and 86.67% respectively [12].

TABLE I. Reported elements on Comparing results from various works

Date	*Work*	*Patients*	*Techniques*	*Achievements*
2006	EXINI v1 [7]	200	threshold+ 14 features + ANN	Sensitivity=0.9 specificity=0.74
2009	EXINI v2 [8]	810	ASM + 45 features + ANN	Sensitivity=0.9 specificity=0.89
2012	BONENAVI v1 [9]	904	ANN	AUC=0.91
2013	BONENAVI v2 [11]	1532	ANN	AUC=0.94 Sensitivity=0.92 specificity=0.84
2016	CADBOSS [12]	60	LSAC+ lPCA (10 features) + ANN	Accuracy =0.92 Sensitivity=0.94, specificity=0.87
2016	LU thesis 1 [13]	2164	CNN	Sensitivity=0.98 specificity=0.65
2017	LU thesis 2 [14]	2164	CNN	AUC=0.94
2020	D.A.E. Larisa [15]	817	CNN	Accuracy =0.97 Recall=0.87 Precision=0.95

Two M.Sc. theses from Lund University used convolutional neural networks (CNNs) for hotspot classification collected from EXINI patients. In the first one, due to execution time restrictions, only the hotspots found in the spine were used to train the CNN. Thus, the used dataset consisted of 10,428 examples where 3,170 of the them belonged to the “high risk” class, while the remaining 7,258 examples were from “low risk”. The dataset was separated into training, validation and test set having 4,171; 3,128 and 3,129 hotspot respectively [13]. The CNN used the Keras tool and under sampling. The second used to find optimal hyper-parameters 20% of the data for the test set to calculate

the final performance, and the remaining divided into equally sized 'folds' for K-fold cross-validation with the number of folds 3, i.e. K = 3) [14]. For all results obtained the AUC was used as a performance index. The final test set performance was calculated by taking the optimum hyper-parameters obtained after a random hyper-parameter search task and retraining a network using the training and validation data combined, thereby extending the number of training examples. Table 1 presents the score of both [14].

Recently, a study investigates the application of CNN to build a computerized solution that automatically identifies whether a patient has bone metastasis or not. In such work, 970 exams of whole body scintigraphy images from 817 different male patients who visited Nuclear Medicine Department of Diagnostic Medical Center (Larisa, Greece) in order to have diagnostic were used [15]. They proposed a CNN model after comparison of a diverse well know models like ResNet50, VGG16. GoogleNET, Xception and MobileNet [15]. The method includes four steps: data preprocessing, feature computations, training phase and validation. The data split saves 15% of the total dataset for testing. The remaining 85% of the dataset is then split into an 80/20 ratio, where the small portion is used as validation set [15]. The proposed model has a deep-layer with 3 convolutional pooling layers, 1 dense layer followed by a dropout layer, as well as a final output layer with one node. The final output was recall 98%, accuracy 97% and precision 95% [15].

By these works, summarized in Table 1, it is possible to state that the developed of bone scan aiding tools has had a advances in the last years, but there are still opportunities for new approaches and specially to promote better interaction between the computer and medical user.

4. IMPLEMENTED SYSTEM

The source code of this project is open source, free and accessible at the following online repositories: github.com/josemorista/bm-server (Back-end) and github.com/josemorista/bm-web (Front-end). The implemented web application is also available online for public access at bm-diag.org. In order to achieve bone metastasis classification and to compute the number pixels labeled as with metastasis, the system performs in the back-end image analysis (IA) and machine learning (ML) operations that will be described in details in next sub sections of this work.

A. *Used Technologies*

The application's interface uses the World Wide Web (WWW) and was implemented using ReactJs, Hyper Text Markup Language (HTML), Cascading Style Sheets (CSS), Syntactically Awesome Style Sheets (SASS) and TypeScript (TS). Available languages for descriptions and communication between the machine and the user in the system interface are English and Portuguese. This front-end application communicates through requests under the Secure Hypertext Transfer Protocol (HTTPS) with an application programming interface (API). The body of the requests is written in JavaScript Object Notation (JSON) and the methods available for communication follow the Representational State Transfer (REST) architecture. The API, mostly implemented with NodeJs, TS, ExpressJs and Python is responsible for handling and processing the various requests entering from the web interface, regarding authentication, management and processing of exams and patients. To perform operations related to IA and ML, the API executes sub processes (or child processes) written in Python language with help of the popular tools: Pydicom; Skimage; Opencv; Scipy; Sklearn and Pandas. Pydicom was used for reading and processing files in DICOM format. The Skimage, Opencv and Scipy are employed to execute the set of operations related to the IA and image processing (IP). Sklearn and Pandas are used for the tasks of classification and learning. The application's back-end also communicates with a PostgreSQL database cluster in order to provide persistence of information related to users, patients and exams.

B. *System security*

Security and access control are managed through a Bearer Token and JSON Web Token (JWT) strategy. The user's JWT is created and signed by the API at the time of the user authentication. This token has a limited lifetime of 30 minutes, being extensible through security routes. This JWT contains the professional's unique identifier in the system and must be attached to the HTTP "Authorization" header in all private route requests. Through this security mechanism, it is possible to guarantee that each professional will be able to view and manipulate data referring only to their profile and patients. These include: upload DICOM files, create, edit and view patient information; create, edit and process exams. Once authorized, the user (physician) can use functionalities to create, edit, list, remove patients and process their exams for automated detection and diagnosis of metastasis. In order to add a patient into the system, one has to include: Patient's full name; date of birth; gender and complementary observations about the patient. Once a patient was added (or selected), the user is redirect to the screen related to the exams of that patient. As presented in Fig 2 already processed exams are assigning with a green dot and exams do not yet processed with a yellow one.

Fig. 2. System screen related to exams visualization and management.

C. *Used dataset description*

The used set of bone scan exams was acquired in the UFF University Hospital (UH). The acquisition was done according to the Brazilian Society of Nuclear Medicine protocol for this exam. The exams must be in DICOM format. The scanner used to perform the exam was a GE Millennium® MPR Gamma Camera. There are 42 available exams, they are 21 on ventral and 21 on dorsal positions. From these, 8 patients have diagnoses of bone metastasis while the remainder 13 patients have healthy radio tracer distributions.

In order to effectively train the classifiers and evaluate the output, a ground truth (GT) of the exams was obtained through medical consensus of the experts from our UH. These experts were responsible for the task of manually identify pixels as belonging to a metastatic area of the images. Used GT were composed of 42 annotated images from the 42 available bone scan exams. The results of this annotations were digital images with metastatic regions segmented and labeled as foreground pixels. To do this operation, the physicians used the web application APEER. Available at https://www.apeer.com/annotate. This online tool provides an easy interface for accurate annotation [16].

D. Approach assumptions

The objective of this work is producing an application for classification of each pixel of the exam as metastatic or non-metastatic without user interaction. In other words, the idea is the design of an application with adequate default options. For it each foreground pixel in the bone scan DICOM file pixel data will be considered a record of the learning database and will have its own feature vector. Each pixel in this exam can range from 0 to 4,000. The pixel is considered foreground when its intensity in the original image is not equal to zero. This means that all pixel with null intensity values in the original image will not be part of the learning dataset, even if by applying some IA filters this positions receive values different of zero. It is worth note that this strategy presents at least two advantages: it provides a larger dataset to be trained by the classifiers, and suits for the goal of provide metastasis classification at a minimum level (that would be extremely difficult for the human eye), that is examining at pixel level.

The implemented strategy to collect the features will be compose the vectors with elements coming from the original intensities, positional attributes and other several features obtained by the application of techniques from IP. Note that these IP techniques are commonly used for image enhancement, edge detection, restoration and others. In the here used strategy, the results produced by these techniques will be used as features and do not have their results visualized or used by next IP steps as usually done in non-autonomous IA and Pattern Recognition (PR) applications [17].

E. Feature vector composition

Initially, for each pixel of the patient body of an exam 65 elements are included in the feature vector. They are formed by integer or Boolean elements. For each exam, the operations of reading and converting the DICOM file in such metadata were executed. After the normalization feature vectors are ready for identification. That is each originally non-null pixel (belonging to the foreground) presents in the exam has its feature vector computed. The data set is transformed to an array formed by 2,512,386 lines (its pixels) and 65 columns corresponding to (its features). These columns are composed by: 1 integer for the intensity of absorption; 2 integers for vertical and horizontal pixel position in the image; 1 binary corresponding to be or not in the pelvic region; 3 integers from Gaussian filters with σ= 3, 5 and 7; 2 integers from Median filters of a 3x3 and 5x5 neighborhood; 1 integer from global histogram equalization; 1 integer from CLAHE operation; 1 binary from Otsu threshold; 1 integer from Roberts filter; 1 integer from Sobel filter; 1 integer from Entropy computation; 1 integer from variance; 48 integers from Gabor filters [17]; 1 binary attribute representing metastasis (1) or do not (0) in the ground truth. This last position of the feature vector is a tag. It is included in order to simplify the notation on the learning phase. After when new exam are used for identification 64 elements of the feature vector of each pixel are computed initially and final one is the definition of being of not a metastatic pixel. That is, the diagnosis with 1 for a metastasis diagnosis and 0 for non-metastatic. Regarding the class distribution of the dataset used for ML 2,457,703 entries belong to the healthy bone class and 54,683 belong to the metastasis class. This unbalanced class distribution is expected since there are very few cases where the affected area in the bone scintigraphy is greater than the normal bone pixels count.

Three of these features need some additional words. For position 4, that is the flag about pelvic region, a template match approach with an average appearance of this region in BS is used for identification of this region in an examination. Figure 3 shows the template *R* used in a search image *I* in order to detect the best matching. The template *R* is positioned from left-to-right and top-to-bottom across *I*. At each (*x, y*), a correlation is calculated to represent the degree of match [18].

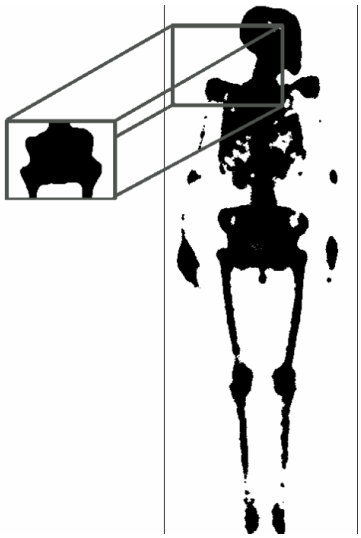

Fig. 3. Template matching operation on processed images

Position 15 of the feature vectors, is the Entropy feature based on the function entropy available in the Python library Skimage. In IP entropy is defined by:

$$Entropy(V) = -\sum P(V)\log_2(V)$$

where *P(V)* is the degree of randomness of a given set of pixels. A local entropy is related to the definition of a certain neighborhood (a rectangular or a circular region), *V* can be considered the gray level of each pixel in the position, *(x, y),* and the probability *P(V)* is the number of occurrence of this gray level in the neighborhood divided by the number of pixels of the region [19]. The configurable parameters of this function correspond to the choice of the shape and size of the region where the entropy will be calculated. The image characteristics used by the application of this function on a circular region of radius 3 around the

pixel. A sample of this filter in a BS is presented at Figure 4. As this filter considerably reduces the scale of the intensities range, due to the log characteristics, in order to visualize the image result, in this work, the result of the floating points values were rescaled to the 0-255 range by applying a linear transformation, using the maximum value computed.

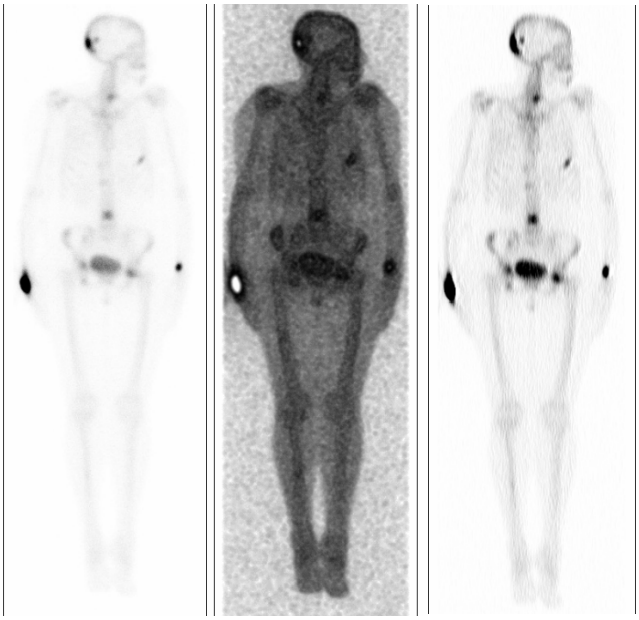

Fig. 4. Original BS image, it after Entropy and Gabor filtering

For computation of the 48 features related to the Gabor filtering, the implementation uses this filter as available at the OpenCv Python library. However, firstly a study was made for determining the adequate range for Gabor parameters to highlight possible hotspot regions. The gamma parameter, referring to the aspect ratio of the Gaussian kernel to be used receives values of 0.01 (almost rectilinear) and 1 (circular or $\sigma_x = \sigma_y$). The study decided on use of 3 angles settled to the θ parameter: 0, π/4 and π/2. The main standard deviation parameter, σ, were also varied to be equal to 1 or 5. The kernel size used was tested by using values of 3 and 7 and the phase offset parameter ψ with values 0 and 1. As for the λ parameter, the tests indicated that lower values than π/4 in the Gabor executions end up overlooking small features at the image. Thus, this parameter was fixed in this value for all filter executions. The left image of Fig. 4 shows results of Gabor filtering with kernels 7×7 and parameters: $\theta = 0°$, $\lambda = \pi/4$, $\sigma = 1$, $\gamma = 0.01$, $\psi = 1$.

Before using these data for training or testing, a normalization of the data is carried out, so that they are transformed to values among 0 and 1, in order to put all of them in a common scale, and do not introduce bias due to possible great differences in their values. To normalize the first feature, that is the intensity of absorption, the minimum and maximum values were considered to be 0 and 4,000 respectively. To normalize the pixel position features, image resolution is considered the maximum (1023 to height and 255 to width) and the minimum value used was 0. For other features the limits were collected directly from the list of computed values on each one.

F. Feature Reduction and Training:

The PCA technique was used to reduce the number of features. The PCA applied was the one available in the Python library Sklearn. The most important 25 components of the feature vector are those that remain. In order to decide about these the results in terms of Explained variance were observed. It is noted that after 25 features, no gained by including additional features is achieved; resulting 25 features and 1 target attribute (i.e. dataset composed of 2,512,386 records).

Next step consists in training and evaluating classifiers. In order to verify the performance of the learning techniques, the train-test split strategy used was K-Fold Cross Validation with a parameter of K = 5. Thus, 5 sets of size 502,477 were generated, each one alternating between training and testing roles. The Gaussian Naive Bayes (GNB), Support Vector Machine (SVM) and Multilayer Perceptron (MLP) with 2 levels were used for performance verification and possible use in the final version.

The implementation used for the Gaussian Naive Bayes (GNB) classifier was the one provided by the function GaussianNB of Sklearn Python library. There were no parameters to be adjusted in this function.

For SVM training, the implementation provided by the Sklearn Python library was applied. No iteration limits were used in the executions in order to enable a longer learning period and without premature interruptions. The Kernel trick function used was RBF, defined by the equation:

$$RBF(V) = e^{(-\gamma p(||V - V`||)2)}$$

where V and $V`$ are entries from a subset of the training data and γp is a positive value defined using:

$$\gamma p = 1/(25 \times \sigma^2)$$

where square sigma (σ^2) represents the variance of a random subset of the input set and *25* is the number of attributes (used after PCA selection).

To evaluate the capacities of the Multi Layer Perceptron (MLP) technique in this problem, it was tested with 10 and 100 hidden layers. The implementation provided by the Sklearn Python library and the selected activation function ReLU was used.

The achieved results are presented in Table II. The MLP 10 is defined for the final version of a fully automatic system because this presents the best values for all evaluators computed.

TABLE II. CROSS VALIDATION RESULTS

Technique	*Recall %*	*Precision %*	*AUC*
Gaussian Naive Bayes	90.45	61.22	0.97
SVM	75.03	83.48	0.97
MLP 10	78.85	84.07	0.98
MLP 100	76.45	83.26	0.98

G. Implementation output:

At the end of the processing the metastatic areas detected by the application can be presented to the user in various ways. The application can produce a report

containing the quantitative evaluations of the findings. The user can also alternate (using the radio button on the right of the screen) between white foreground and black background or a black foreground and white background.

5. Conclusions and Final Remarks

Comparing last column of Table I (discussed in section 3) with Table II, the AUCs of this implementation present better results while for Recall and Precision the results are not so good then those of related works. However, the number of patients used in this 16 years on various works can allow us to state that the ideas presented in this application to deal with a very reduced GT (that is the considerably lower number of patients and samples used) enable to achieve an efficient tool for real time uses. Presented ideas contributed to build a solid application increasing the number of samples available for training and testing the hotspots and, reducing their size to a minimum.

Nowadays, a number of works can be found applying artificial neural networks (ANN) and convolutional neural networks (CNN) to help in BS diagnosis [7 - 15]. Although, especially in this imaging such a number is not so huge as in the other medical aspects, where it is almost impossible to find recent works (in height impact journal or quality conferences) that do not use one of these techniques of Artificial Intelligence (AI) and Machine Learning (ML) [20, 21]. However, in Computer Science as other areas of research, diversity is fundamental for good science making. In this direction, the here presented and developed solution allows AI/ML tasks to depend only of simpler image processing and decision techniques to detect and measure such bone metastasis. Moreover, all the presented developments, even being traditional techniques have here a new programming approach. The main aspect of this is the way the features are computed from the images and used in the decision process.

There are 2 original ideas in the work here presented. The first focus in the way normally enhancement or the also named preprocess step is considered in Digital Signal and Image Processing. The second is related to the image area from where the feature are extracted in the analysis.

In the related implementation most of the used features are normally employed only for image preprocessing. This is the case of the techniques of Histogram Equalization, Contrast Limited Adaptive Histogram Equalization (CLAHE), Otsu thresholding, Median, Gaussian and Entropy filtering, Roberts and Sobel edge detection, Variance computation and Gabor filter banks. Traditionally, all of them are used to transform an image (or a part of an image) in a enhance one related to some concepts [17]. However, they are here converted in an numeric element of the feature vector commented in sub section *E*. The knowledge of their numbers after normalization in each pixel related to the patient body are proposed to be features for the ML techniques.

The second original aspect of this work is related to the area used for computation of the feature vectors to be used to detect anomalies and final decision about each one is (due the amount of gamma ray emission) with or without metastasis. This can be seen as related to region of interest (ROI) or segmented areas in traditional pattern recognition implementations [17]. As mentioned in subsection E, we presented the idea of use all non background pixels in the image on analysis as ROI candidates to be analyzed. For each one, all features are computed and the feature vector analyzed. This allows having as many cases as possible to learn even in a very small set. As far as we know such consideration of minimum possible area from analysis or maximum possible set of feature vectors in a study was never presented and reported before. Of course, the incipient number of patients that are available to us for this development has directed to this use. However, it is important to say that the existence of so few cases to use can be the case present in many other circumstances and also represents important peculiarities that must be consider in other cases and applications. Maybe, it is paramount to realize that there will be always new issues to study and possible aspect that present few samples for learning solutions. In these cases a possible solution for the problem is better than no solution at all. New diseases could appear and spread quicker than the capacity of the medical community to organize annotated datasets to be used in ML supervised approaches or to allow CNN and ANN uses. The knowledge of all possible techniques is important not only for historical and academic point of view but mainly to allow the development of new solutions in any circumstance. This and their uses are very important for real cases in unpredictable future and environments.

Successful development of new and real applications involves proper uses of both traditional and artificial intelligence concepts. Complete knowledge of their basic characteristics is needed to decide when each one is more relevant to a specific aspect of a problem on focus. Considering this, there is an obvious hope to be able to overcome, in a near future, the challenge of having more data for this application. Then with more annotated dataset and with techniques such as data augmentation and transfer learning, the techniques of ML and data mining can be used with more efficiency in this so important aspect of nuclear medicine [2 - 15]. So as future works in this specific subject, it will be possible to address the steps of training with new learning algorithms. Also, some improvements can be made regarding the experimentation of modern classifiers such as XGBoost and the use of an auto-learning strategy, in order to delegate the task of selecting the ML algorithm that best fits to an automated solution. Regarding the pre-processing stages of training, strategies such as data augmentation and under sampling of the true negatives class can also contribute to the improvement of the results in next works. Train-test split strategies other than K-Fold cross validation can also be tested as future work, as an example; one common approach in medical ML problems is the use of a technique called leave one patient out. This strategy works analogously to cross-validation, however, the selected test set at a time corresponds to a subset of data belonging to a single person. This technique helps to isolate from the training set information with a very close level of similarity (same patient), which in some cases can bias the obtained results.

The concrete goal or at least one of the more important aims of this study is to help medical doctors to analyze the disease evolution by scintigraphy exams and judge if the bone metastasis is growing or stable. However, for the here

presented work, only direct measurements was discussed. So, a way to automatically evaluate such modifications could be compare the area of pixels associate with the metastasis body points after Register exams of same patient in two dates [22]. (In image analysis, Register is a technique that, basically, takes two images and transforms one image into the other, allowing match objects common in both, so that similar points in world space are also similar in image coordinate space.) After this match patient evolution can be fairly compared using same techniques. The disease progress evaluation then can be properly made by subtract the images. This is a simple and direct improvement for future works.

The usability of the system was evaluated in the research that originated this work [23]. The Bone Scan Index (BSI) can also be added in the application as another quantitative measure [2]. In order to achieve such measure, the already available functionality of calculating the ratio of metastasis over foreground pixels must be extended in order to be calculated only considering pixels that belong to the patient skeleton. Segmentation of the skeleton in a bone scan image is a task that requires a set of well-defined IA and IP operations and ground truth images in order to evaluate the accuracy of the segmentation.

Another important incoming aspect for the application improvement is the support for processing bone scan from small regions and not only whole-body exams. With this, physicians will be able to use all the already available functionalities (to whole-body) also to process localized and small regions scintigraphy exams such as scans containing only cranium, vertebral column and others. In order to do this small regions bone scans classification, some tests must be made, including training new classifiers with some adjusted features. Although the majority of the implemented features applied at full body scans can be easily used in small regions, some features like pixel position and pixel on pelvic region needs to be adapted.

ACKNOWLEDGMENT

The authors thank CAPES. A.C. thanks to CYTED Spanish agency, FAPERJ and CNPq Brazilian agencies for supporting received in relation to this work.

REFERENCES

[1] Adellia A., How Healthcare Industry Will Change in 2022. HashMicro. https://www.hashmicro.com/blog/how-healthcare-\industry-will-change-in-2022/

[2] Wyant, T. Bone metastases. Cancer.org. 2021 Available at: https:// www.cancer.org/treatment/understanding-your-diagnosis/advanced-cancer/ bone-metastases.html. Accessed: at 28 July 2022.

[3] Wyngaert, T.; Strobel, K.; Kampen, W.; Kuwert, T.; Bruggen, W.; Mohan, H.; Gnanasegaran, G.; Delgado-Bolton, R.; Weber, W.; Beheshti, W.; Langsteger, M.; Giammarile, W.; Mottaghy, F.; Paycha, F. The EANM practice guidelines for bone scintigraphy. European Journal of Nuclear Medicine and Molecular Imaging. 2016

[4] Rieden, K. Conventional imaging and computerized tomography in diagnosis of skeletal metastases Radiologe. Der Radiologe. 1995.

[5] Westmead BCI Institute. Bone Metastasis. Westmead BCI Institute. 2017. Available at: https://www.bci.org.au/breast-cancer-information/fact-sheets/ bone-metastases/. Accessed at: 20 July 2022.

[6] Łukaszewski, B.; Nazar, J.; Goch, M.; Łukaszewska, M.; Stępiński, A.; Jurczyk, M. Diagnostic methods for detection of bone metastases. Contemporary oncology (Poznan, Poland), 21(2), 98–103. https://doi.org/10.5114/wo.2017.68617. 2017.

[7] Donohoe K.; Brown M.; Collier, D.; Carretta, R.; Henkin, R.; O'mara, R.; Royal, H. Society of Nuclear Medicine Procedure Guideline for Bone Scintigraphy. 2023.

[8] Hopkins Medicine. Bone Scan. Hopkins Medicine. 2021. Available at: https://www.hopkinsmedicine.org/health/treatment-tests-and-therapies/ bone-scan?amp=true. Accessed at: 10 July. 2022

[9] Erdi, Y.E.; Humm, J.L.; Imbriaco, M.; Yeung, H.; Larson, S.M. Quantitative bone metastases analysis based on image segmentation. J. Nucl. Med. 38: 1401–6. 1997

[10] Ulmert, D.; Kaboteh, R.; Fox J.J. A novel automated platform for quantifying the extent of skeletal tumour involvement in prostate cancer patients using the Bone Scan Index. Eur Urol. 62(1):78-84.PubMedGoogle ScholarCrossref. 2012.

[7] Sadik, M. Jakobsson, D.; Olofsson, F.; Ohlsson, M.; Suurkula, M.; Edenbrandt, L. A new computer-based decision-support system for the interpreta┐ tion of bone scans. Nuclear medicine communications. 2006.

[8] Sadik, M. Hamadeh, I.; Nordblom, P.; Suurkula, M.; Hoglund, P.; Ogl┐ Soon, M.;Edenbrandt, L. Computer-Assisted Interpretation of Planar Whole-Body Bone Scans. J Nucl Med. 2008.

[9] Nakajima, K; Edenbrandt L; Mizokami, A. Bone scan index: A new biomarker of bone metastasis in patients with prostate cancer.Int J Urol. 2017.

[10] Horikoshi, H.; Kikuchi A.; Onoguchi M. Computeraided diagnosis system for bone scintigrams from Japanese patients:importance of training database. Ann Nuclear Medicine. 2012

[11] Nakajima, K.; Nakajima, Y.; Horikoshi, H. Enhanced diagnostic accuracy for quantitative bone scan using an artificial neural network system: a Japanese multi-center database project. EJNMMI Res. 2013.

[12] Aslantas, A.; Dandil, E.; Saglam, S.; Çakiroglu, M. Cadboss: A computer-aided diagnosis system for whole-body bone scintigraphy scans. Journal of cancer research and therapeutics, 12(2), 787. 2016.

[13] Dang, J. Classification in bone scintigraphy images using convolutional neural networks. Master's Theses in Mathematical Sciences. Lund University, Sweden, 2016.

[14] Belcher, L. Convolutional neural networks for classification of prostate cancer metastases using bone scan images. Theoretical Physics Masters Project, Lund University. Sweden , 2017

[15] Papandrianos N.; Papageorgiou E.; Anagnostis A.; Papageorgiou K. Bone metastasis classification using whole body images from prostate cancer patients based on convolutional neural networks application. PLoS ONE. 2020.

[16] APEER, An intuitive annotation tool for your machine learning needs. APEER. Available at: https://www.apeer.com/annotate. Accessed at: 30 dec. 2021.

[17] M. Sonka, V. Hlavac and R. Boyle, Image Processing, Analysis and Machine Vision, 3th Edition, Thomson, 2008.

[18] Rosebrock, A. OpenCV Template Matching. Pyimagesearch. 2021 Available at: https://www.pyimagesearch.com/2021/03/22/ opencv-template-matching-cv2-matchtemplate. Accessed at: 20 July. 2022.

[19] Scikit image. Entropy. Scikit image. 2021. Available at: https://scikit-image. org/docs/dev/auto_examples/filters/plot_entropy.html. Accessed at: 20 July 2022.

[20] Litjens, G., Kooi, T., Bejnordi, B.E., Setio, A.A.A., Ciompi, F., Ghafoorian, M., van der Laak, J.A., Van Ginneken, B., Sánchez, C.I.: A survey on deep learning in medical image analysis. Med. Image Anal. 42, 60–88 (2017).

[21] Haskins, G., Kruger, U. & Yan, P. Deep learning in medical image registration: a survey. Machine Vision and Applications 31, 8 (2020). https://doi.org/10.1007/s00138-020-01060-x

[22] Myronenko, A., Song, X.: Intensity-based image registration by minimizing residual complexity. IEEE Trans. Med. Imaging 29(11), 1882–1891 (2010).

[23] Carneiro da Silva, J. M., Bone-scan: a Computer-aided Detection (CADe) and Diagnosis (CADx) System for Bone Scintigraphy, Compter Graduation Program, Master's Dissertation, Computer Institute, UFF, 2022 http://www.ic.uff.br/index.php/pt/pos-graduacao/teses-e-dissertacoes