RESEARCH ARTICLE

Evaluation of the Relationship Between the Age of the Patient and the Mandibular Trabecular Bone Structure Through Dental Cone Beam Tomographic Images by Means of a Convolutional Neural Network

Avaliação Da Relação Entre a Idade Do Paciente E a Estrutura Óssea Trabecular Da Mandíbula Através Da análise De Imagens Tomográficas Dentárias Cone Beam Por Meio De Uma Rede Neural Convolucional

Matheus Raffael Simon[1]*, Adair Santa Catarina[2], Adriane Yaeko Togashi[3]

Abstract: Osteoporosis is a systemic condition that affects bone mineral density in individuals, affecting mainly women and is diagnosed by DXA examination. The research used 137 cone beam tomography scans, from which 1389 mandibles samples from female patients and 633 from male patients were selected for CNN training, which is composed of 3 dense layers with 100 neurons each and Relu activation function, update of weights by the Adam algorithm, using MaxPooling in each convolution; the dense layer uses 100 neurons, with the activation function. The neural network training accuracy was 98% for males and 94% for females, with area under the ROC curve (AUC) equal to 0.94 and 0.81 for the respective genders. The accuracy obtained in validating the CNN was 98% for males and 89% for females. In the supervised test, using 5 cuts from 10 exams of the test set of each sex; approximately 100% accuracy was obtained for both sexes. Thus, it is concluded that the proposed model is capable of classifying the samples in the proposed age groups and proved to be robust and solid in the tests.

Keywords: Imaging processing — convolutional neural network — trabecular bone structure — mandible — cone beam tomography — osteoporosis

Resumo: A osteoporose é uma condição sistêmica que afeta a densidade mineral óssea em indivíduos, atingindo principalmente as mulheres e é diagnosticada pelo exame DXA. A pesquisa utilizou 137 tomografias cone beam, das quais foram selecionadas 1389 amostras da mandíbula de pacientes do sexo feminino e 633 para o sexo masculino para treinamento da CNN, que é composta por 3 camadas densas com 100 neurônios cada e função de ativação Relu, atualização de pesos pelo algoritmo Adam, utilizando MaxPooling em cada convolução; a camada densa utiliza 100 neurônios, com a função de ativação. A acurácia de treinamento da rede neural foi de 98% para o sexo masculino e 94% para o feminino, com área sob a curva ROC (AUC) igual a 0.94 e 0.81 para os respectivos sexos. A acurácia obtida na validação da CNN foi de 98% para a sexo masculino e 89% para o feminino. No teste supervisionado, utilizando 5 cortes oriundos de 10 exames do conjunto de teste de cada sexo; obteve-se aproximadamente 100% de acurácia para ambos os sexos. Assim, conclui-se que o modelo proposto é capaz de classificar as amostras nas faixas etárias propostas e se mostrou robusto e sólido nos testes.

Palavras-Chave: Processamento de imagem — rede neural convolucional — estrutura óssea trabecular — mandíbula — tomografia cone beam — osteoporose

1. Introduction

Bone density loss is a metabolic/physiological condition, characterized by the progressive diminishment in bone density leading to an increased risk of fractures, known as osteoporosis. Losses of 10% to 15% characterize osteopenia, a state prior to osteoporosis and which prompts physicians to take

preventive measures [1]. The detection of osteoporosis is usually made through the densitometry exam (DXA – Dual-Energy X-ray Absorptiometry). The DXA is the gold standard test for density analysis [2], but it presents limitations to a large share of the population as, in the Brazilian public health network, there are but few devices – and these are poorly distributed among the five geographic regions of Brazil.

The bone from the human skeleton is basically divided into two parts to the naked eye: the trabecular bone, which presents a spongy aspect; and the compact bone, which is denser. Bone mass changes throughout life and, around age 35, it reaches its peak. From this age onwards, a slow process of bone loss begins, getting more accelerated in women during menopause [3].

Cooper and Melton III [4] built a timeline for bone mass across the years of life, shown in Figure 1. As it can be seen in Figure 1, (♂) represents the timeline for males, (♀) represents the timeline for females. (I) corresponds to the peak of bone mass achieved at approximately 35 years of age for both genders. (II) is the drop in bone mass during menopause in women. (III) is age-related bone loss for both genders. The horizontal axis represents age in years, while the vertical axis represents the amount of calcium (in g) stored in the skeleton of an individual. The density of the trabecular part of hu-
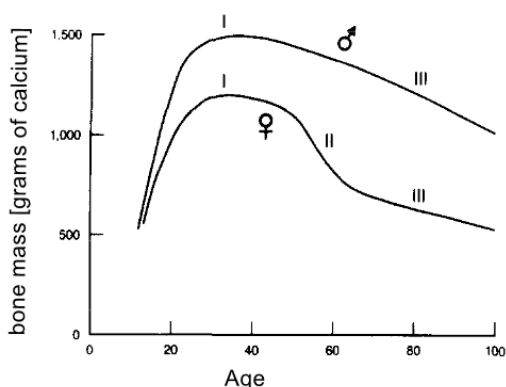

Figure 1. Men and women bone mass timeline [4]

man bone decreases over time after reaching its peak, and in women the density loss is greater due to menopause [5][6][7]. The study by Von Wowern, Klausen and Kollerup [8] reported that osteoporotic individuals had lower mandibular bone mineral content when compared to normal individuals. Dentists need to know the anatomical details of mandibles of their patients in order to place dental implants. Cone-beam computed tomography (CBCT) is usually used to analyze said region because it provides a three-dimensional image of the face of the patient, presented in three anatomical planes – axial, coronal, and sagittal –, registering values on a standard scale, called Hounsfield Units (HU). In this way, means can be used to find patterns in images, such as, for example, the use of a neural network.

The use of neural networks, artificial intelligence, and machine learning is increasingly present in the healthcare area, assisting in the diagnosis of medical images. Especially in the classification of images that present some type of disease [9]. Convolutional neural networks are used to find patterns in images by applying filters on visual data [10].

Based on the premise that the DXA scan is not accessible to a significant part of the Brazilian population and that there is an association between bone trabecular structures and bone mineral density, this study aimed to develop a methodology that uses a CNN to analyze the relationship between the age of a patient and the trabecular bone structure of their mandible through CBCT. The CNN classifies patients into different bone age groups according to their gender.

Therefore, patients whose ages are not compatible with their bone ages can be referred for an assessment of their bone mineral density. The method developed and coded in software does not diagnose osteoporosis, what it does is subsidize the dentist to refer the patient to a doctor who will eventually request a DXA scan, which optimizes the use of equipment and resources of the Unified Health System (SUS).

1.1 Bone Structure of the Mandible and Osteoporosis

Bone tissue is a specialized form of connective tissue comprised of cells and an extracellular matrix, which has the unique characteristic of mineralizing. The mineralization of the matrix provides this tissue with an extreme stiffness, allowing it to perform important functions of support and protection [11].

The jaw is a mobile bone of the face. It is shaped like a horseshoe; its central region is called the body of the mandible, while the lateral areas are the ramus of the mandible. The mandible is a unique bone, located in the lower third of the face, articulating with the glenoid cavity of the temporal bone, where it forms the temporomandibular joint (TMJ).

The muscles attached to this bone are responsible for swallowing, phonation, mastication, and facial expression [12].

Osteoporosis can be characterized by reduced bone mass, thus leading to bone fragility and susceptibility to fractures. This disease can be confined to just one limb or be widespread in the entire skeleton. Primary forms of osteoporosis are more common and may be associated with aging or postmenopausal status in women [13].

Osteopenia is the decrease in bone mass, caused by the loss of calcium and vitamin D deficiency, which may lead to osteoporosis; in other words, osteopenia is an earlier stage that signals a possible evolution to osteoporosis [14]. Osteoporosis is often closely linked to aging, but some other risk factors contribute significantly to the onset of the disease such as: natural degeneration, calcium deficiency, lack of vitamin D, smoking, alcoholism, high caffeine intake, menopause, family history, physical inactivity, and some thyroid disorders [15].

Osteoporosis is a critical disease and can reduce both life expectancy and quality of life in individuals with such disease,

but the likelihood of developing it can be reduced with healthy habits such as physical exercise, sun exposure, intake of food rich in calcium, and avoiding alcohol, tobacco, and caffeine abuse [16].

Luize et al. [17] suggests in their study that the maxilla and mandible, in addition to long bones and vertebrae, may present alterations related to osteoporosis, and Von Wowern, Klausen and Kollerup [8] mentions that the mineral density of the mandible of people with osteoporosis is lower than the bone density of people without this disease. Thus, it is understood that the mandible may present a difference in its bone structure between the bone mineral density of a healthy mandible and one with osteoporosis.

1.2 Convolutional Neural Network

Convolutional neural networks (CNN) are a special type of artificial neural networks (ANN) used in processing data that are organized in mesh shapes with one or more dimensions [18]. Neural networks are based on the functioning of a biological neuron. In essence, a convolutional neural network is formed by three layers: the convolution layer, which is the layer where the images are converted into data processable by kernels (filters) and the main features are returned; the pooling layer, which is the layer that simplifies the information from the previous layer through methods such as maxpooling, average pooling, and L2 pooling; and the fully-connected layer, which is responsible for processing and assigning a probability for classification into an output class [18].

CNNs are designed to process data that come in the form of multiple matrices, such as images, which are subjected to filters and stages where their characteristics are removed and saved, resulting in non-linear data [19]. In a simplified manner, the architecture of a CNN is based on the alternation of convolution layers, which gives the network its name, and pooling layers [20].

1.3 Specific and General Objectives

As for the specific objective, it consisted of developing a software to analyze the CBCT images selected by the dentist and to classify the patient into one of the training age groups. If the age group resulting from the analysis is different from the actual age of the patient, that will be an indication of alteration in the trabecular structures of their mandible.

In these cases, the dentist may refer the patient for an appointment with a doctor who will evaluate the bone health conditions of said patient.

The general objective of this study is the development of a new methodology that uses a CNN to evaluate and classify the bone structure of the mandible of patients into age groups, separating them by gender, based on the analysis of multiple layers of a CBCT.

Two databases were created for the training of the CNN: the database for men and the database for women. Table 1 represents the classes and age groups for men, and Table 2 represents the classes and age groups for women. Classes were defined based on the variation in bone mass of human beings

Table 1. Classes and ages for Men

Class	Ages
1	18 – 35
2	36 – 50
3	51 – 70
4	+71

Table 2. Classes and ages for Women

Class	Ages
1	18 – 35
2	36 – 50
3	51 – 60
4	61 – 70
5	+71

as a function of their age, as shown in Figure 1. For women, an additional class (class 3) was established, corresponding to the period of menopause, when there is a sudden decrease in bone mineral density caused by hormonal changes.

1.4 Related Studies

Some studies presenting the same objective were found, such as Togashi et al. [21] that developed an algorithm to process images and quantify the quality of the trabecular bone of the mandible of patients from the analysis of panoramic digital X-ray images, classifying the density into two classes: normal and abnormal. Using digital radiographs images and being accompanied by a dental professional for the assessment of requirements, it was found that the implemented software has the potential to assist dental professionals in the evaluation of the bone structure of their patients and in the referral of patients with suspected low density.

Togashi et al. [21] showed in their work that it is possible to use the trabecular structure of the bone to create a scale to observe bone density.

Lee et al. [22] evaluated the performance of a computer-aided diagnostic system, based on a convolutional neural network, for the detection of osteoporosis in panoramic radiographs, through a comparison with diagnoses made by oral and maxillofacial radiologists. In the study, 1268 panoramic radiographs of women were used, with an average age of 52.5 years, with a standard deviation of 22.3 years, divided into three groups: normal; low erosion; severe erosion, in this group osteoporosis was confirmed. The accuracy results presented by the authors reached 99%.

Lee et al. [22] used a CNN model that presents good accuracy, even with a low number of intermediate layers.

Tang et al. [23] designed a method to classify exams by means of CT images using a CNN. The proposal of this study was to use two modules to detect the condition of the lumbar vertebra bone and determine its quality. The first module of this research locates the ROI in the tomography; it uses a CNN, with U-Net architecture, to capture the ROI (lumbar

vertebra) and to segment the image, discarding the other details that are not of interest in the image. The second module is responsible for classifying the bone into three categories, namely: normal; low mass loss; and osteoporosis. This study had an accuracy of 76.65% and an AUC value of 0.91. According to the authors, the system proved to be a good tool in helping radiologists analyze CT scans and provide low bone density alerts.

The research by Tang et al.[23] demonstrates that it is possible to classify a bone by a CNN and that image segmentation is necessary to avoid image noise.

The present study differs from the previous ones in some respects, such as: the use of cone beam tomography, the segmentation and analysis of trabecular structures of the mandibular bone, and the indirect assessment of bone density through the evaluation of the relationship between age and trabecular bone structure of the mandible of patients. The developed CNN classifies trabecular structures into different age groups that, when older than the actual age of the patient, indicate cases in which a medical specialist should be consulted.

2. Materials and Method

2.1 Acquisition of Cone Beam Tomographs

In order to carry out this study, 137 CT scans of the mandible of patients were obtained from the Dental Clinic of the State University of Western Paraná (Unioeste), 85 of which were female patients and 52 were male patients, distributed according to Table 3. This table had the data grouped for the two genders, thus, the classes represented by "-", indicate that this class does not exist for that gender.

The images come from a CT scanner model ORTHOPANTOMOGRAPHTM OP300, manufactured by the company Instrumentarium Dental, PaloDEx Group Oy, from Finland, in 2016. Its field of view (FOV) is of 61X78 mm and the software used in the processing of the acquired images is OnDemond3DTM Dental. The CMOS sensor, which equips the tomograph, has a resolution of $100\mu m$ per pixel.

The images taken from the DICOM format, correspond to the axial plane, have 300 layers and spatial resolution, or size, of 400x400 pixels, and all samples are in grayscale, with a size of 16 bits per pixel.

The study, within its limitations, used exams with certain criteria in the assembly of its database, such as: CT scans of male and female patients; CT scans of patients with radiographic report without the presence of bone pathologies; and CT scans of patients with all teeth, partially and completely edentulous. This database was created in conjunction with Bottega's master's thesis [24], which studied the correlation of local and systemic factors with the bone density of the jaws obtained through cone beam tomography. Therefore, the database does not contain images of patients with bone pathologies.

Thus, the set of cone beam tomographic exams was divided into two databases, one for training and the other for testing, in the fractions 60% and 40% respectively, respecting

Table 3. Distribution of exams by age group

Gender	Classes					
	18-35	36-50	51-60	61-70	51-70	+71
Female	6	16	31	23	-	9
Male	12	15	-	-	21	4

the proportionality of sexes and age groups.

2.2 Software

The software was developed to be used by dentists and other healthcare professionals, as it facilitates the storage and retrieval of information related to the CBCT scans from their patients. When using the software, users enter the necessary information about their patients, their exams, and perform the classification of the age group associated with the bone structure of the mandible from the patient, comparing it to their real age at the time of the CBCT scan.

The software has standardized parameters obtained during the development of this study, but it offers the user the possibility of obtaining them from a new database and, consequently, new CNN training, which can only be conducted by a user authorized by the system administrator. In order to execute the training, the user may also adjust the metrics used in the training, such as the number of images analyzed by the CNN, the number of training epochs and the number of images to be tested. All this information remains stored in a database and is available to the user at any time.

2.2.1 Reading DICOM Images and Conversion to a Hounsfield Scale

The developed software has a method which reads all DICOM files from a tomography, transforming them into an array that contains the intensity value of each pixel of the image. The matrix data are resized to the Hounsfield scale values using an existing metadata in each DICOM, the tag (0028, 1053), also known as the RescaleSlope key. The value of this metadata is specific to each tomograph manufacturer and is used to linearize the pixel value on the Hounfield scale. The process of resizing to values on the Hounsfield scale is performed through the Equation 1, where: HU = Hounsfield Unit Value; VP = Pixel value from the CT image; RS = Value of the RescaleSlope key from the CT scan or the tag (0028, 1053), where: PS = Value of the smoothed pixel; VA = Value of the image pixel; MIN = Value of the lower intensity pixel in the image; MAX = Value of the higher intensity pixel in the image. After all the pixels of each DICOM image are normalized, the initial array has its values in the Hounsfield scale.

$$HU = (VP * RS) + (-1024) \qquad (1)$$

2.2.2 DICOM Data Presentation

Data normalization is done to optimize the reading and display time of all CT slices, preserving the original data saved,
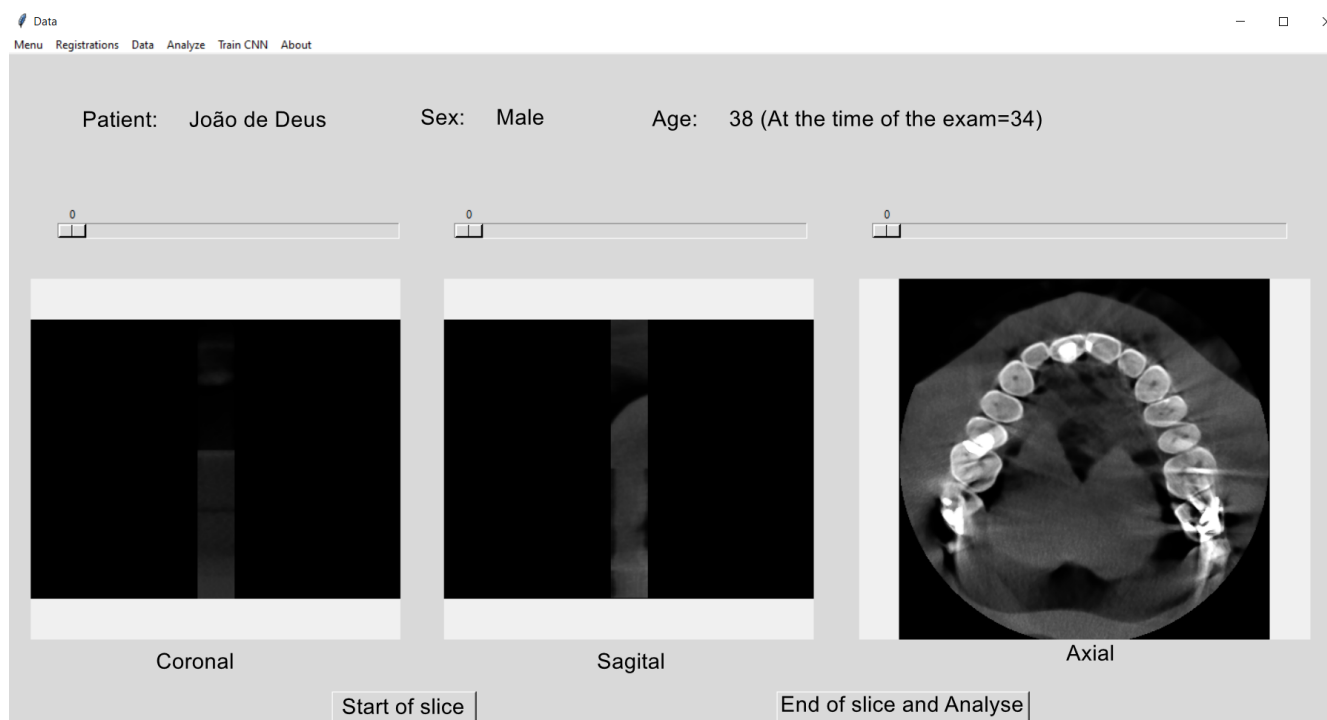
Figure 2. CT analysis screen

in order to accelerate image loading in the graphical user interface. The normalization performed on each pixel of the image follows the Equation 2. Once the normalization of the image is done, the interaction with the elements of the graphical interface became faster and the time needed to load and display the tomographic images was also reduced.

$$PS = ((VA - MIN)/(MAX - MIN)) * 255 \qquad (2)$$

2.2.3 Slice Selection in CT scan

When analyzing a CT scan, much of its imaging is not used, as only specific regions or slices are considered in the analysis. To avoid unnecessary processing and reduce the time consumed, a slicing method was incorporated into the software, thus allowing the user to choose the layers corresponding to the initial and final sections, defining the number of samples that will be analyzed. The plane used here is the Coronal plane, as the images in this plane provided the best results in preliminary tests.

Bone density in the body of the mandible is higher than in the ramus of the mandible [25] and, for this reason, the software uses samples collected from this ROI. The initial slice must be selected after the cortex of the body of the mandible, and the final slice cannot reach the posterior cortex of the body of the mandible. Thus, in this interval, there are only portions of trabecular bone to be processed. The minimum number of selected samples must be equal to five, as the result of the CNN analysis will be gathered through the average of the percentages of the classes of each slice. Figure 2 shows the graphical user interface of the software for selecting the CT slices to be analyzed.

2.2.4 Image Segmentation, Data Extraction and CNN's Input

After the sample selection, the software is redirected to the task of segmenting the trabecular structures, discarding CT images that are outside the range selected by the user. At first, the implemented strategy discards all pixels with values above the dense bone mass, that is, above 800 HU. This was shown to be necessary to prevent noise from interfering with image analysis by the CNN. Another strategy used was to discard all pixels between line 0 and line 65, since it was observed that no valuable information was found in the images from this interval. Therefore, only the region of the images containing the mandible is preserved. Figure 3 (a) shows the partially segmented sample.

Subsequently, all pixels smaller than 300 HU and larger than 800 HU are discarded, in order to keep only the trabecular structures of the mandible. The values 300 HU (minimum) and 800 HU (maximum) were adapted from [26] and [27] because these studies show slightly different Hounsfield values for trabecular bone: between 236.4 HU and 751.1 HU [26], and between 385 HU and 763 HU [27]. These values were obtained from CT images of patients of both genders and of different ages. In order to achieve the expected result, shown in Figure 3 (a), the strategy used to discard pixels consisted of rotating the image, seen as the identification of pixel intensity is performed through a linear repetition structure. Each row of the data matrix is traversed, and its values are nullified until it finds a Hounsfield value corresponding to trabecular bone (800).
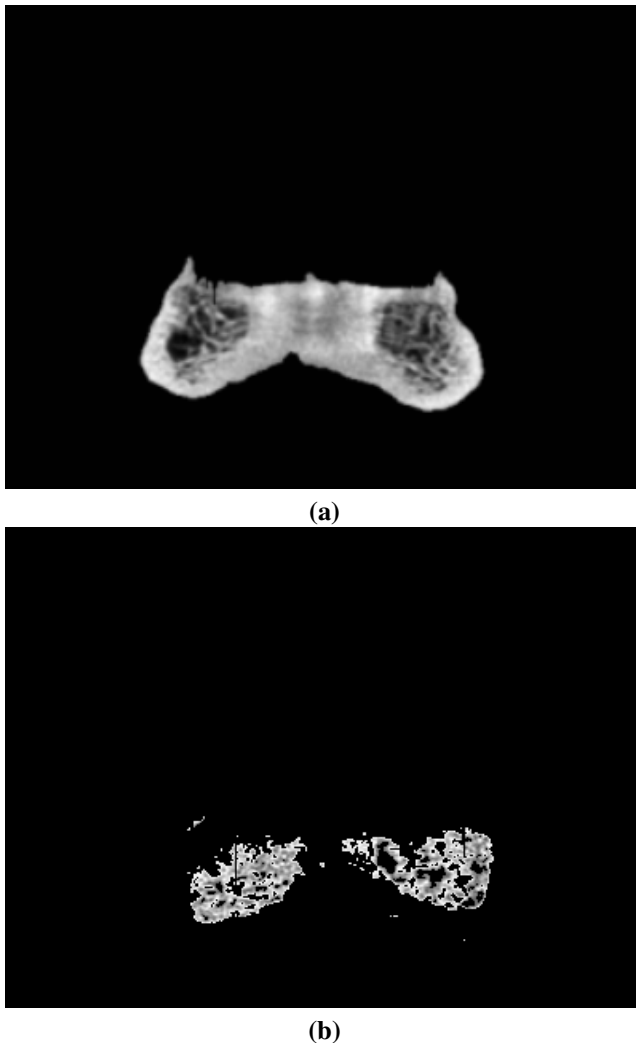
(a)

(b)

Figure 3. Segmentation of the samples

After processing the last row of the image matrix, it is rotated by 90º, and the same process is repeated until the 4 sides of the image have been evaluated. This strategy was necessary because the preliminary tests showed that some regions of the samples were not segmented properly, including noise in the samples. Finally, the sample of the body of the mandible shows only values referring to the trabecular bone and the remainder of the pixels with value zero. Figure 3 (b) displays the result of a fully segmented image. Now that image is CNN's input.

2.2.5 CNN Configuration

After several empirical tests, with different structures, the model considered as the one that provided the best results was reached. This model had great inspirations in the models proposed by the works of Lee et al. [22] and Tang et al. [23]. However, with some differences and adaptations for this study. Some of the hyperparameters, such as the learning rate for example, adopted the default parameter, which has already been extensively evaluated by the authors of the libraries used

in the work.

The training method was created with a convolution operator (kernel) in 2D, utilizing 64 kernels sized (3, 3), that is, a 3x3 matrix. The input image is transformed into a size of 64x64 pixels. The activation function chosen was Relu. After convolution, pooling is applied to extract features from the matrix convoluted by the kernel. The pooling method chosen was MaxPooling of size (2, 2). Finally, flattening is performed, which transforms the data, or characteristics, into a vector.

These were the convolution operators employed. The fully connected neural network was built to be dense, with three hidden layers, each layer with 100 neurons and Relu activation. The output layer is also dense, composed of the same number of classes for the age groups, that is, when training is aimed at males, it will have 4 classes, when females, 5 classes, and it has softmax activation.

The Keras 'ImageDataGenerator' method was inserted to increase the amount of images to be processed by CNN. This method does not create images, it just transforms the original images, in real time and at each epoch, to obtain images with different angles, zoom and rotations, in order to expand the original database. In preliminary tests performed, the accuracy of CNN increased with the use of this method, so it was included in the final version of CNN. It was used to obtain a greater variety of images; the attributes applied were: rescale = 1/255, rotation range = 7, horizontal flip = True, shear range = 0.2, height shift range = 0.07, and zoom range = 0.2. We then move on to training, through the flow from directory method, with the following attributes: directory= 'path of images', target size = (64,64), batch size = 32, class mode = 'binary'. In the model compilation part, the Adam Optimizer Algorithm was used, which calculates the adaptive learning rates individually for different parameters from the estimates of the first and second moments of the gradient. This method also starts the value of the biases in an analogous way, in a second moment this value can be updated. The loss calculation function of the model is the "Sparse Categorical Crossentropy". Also, during the training process, a validation is carried out.

Validation is a process that helps a CNN refine the weights used in training. The Process validation consists of performing, at the end of each epoch, a classification with a set of data different from the training set, with that the CNN adjusts the weights used in the next epoch. At each training epoch, 30 images are randomly chosen for classification. The Keras sequential model was used in the construction of the CNN because this model proved to be more flexible than others. The CNN uses the concept of Backpropagation to update weights during its training.

3. Results and Discussion

3.1 DICOM Image Reading

The proposed software performs the reading of DICOM images from a CT through a method that receives as a parameter

the path to the CT. The method then traverses the directory pointed out in the path, reading and storing in a list only the files with the extension "dcm", which is associated with DICOM type files.

3.2 Transforming Images Read into Axial, Coronal, and Sagittal Planes

A CT scan, by default, is acquired by the CT scanner as a sequence of axial planes. Thus, obtaining the axial planes requires no effort since the DICOM file itself is already created in this way. To obtain the remaining sagittal and coronal planes, it is necessary to obtain the pixel array of each DICOM image of the CT. The pixel array is an element that contains all the values, in bytes, of the pixel intensities. With this data, it is possible to collect and shape the sagittal and coronal planes based on the indices of the pixel array of the axial plane. To model it in another plane, it is necessary to remodel the indices and values of each slice of the DICOM images, adjusting their positioning into a new list.

3.3 Image Normalization

This method was created to accelerate the display of images in the graphical interface of the software. Each axial, sagittal, or coronal list contains approximately 400 images, that is, the graphical interface worked with approximately 1200 images in memory. It was decided to convert the images to grayscale, while also keeping the original images in memory to be worked on later. Equation 2 is used to smooth the images.

3.4 Convolutional Neural Network

3.4.1 Training Method

Several tests were carried out and the training configuration with the best cost-benefit ratio for female patients was performed with 30 training epochs and 300 images, providing training accuracy equal to 94%, consuming approximately 2 hours and presenting an AUC curve equal to 0.94. For males, the choice was to train the CNN with 30 epochs and 100 images, obtaining an accuracy equal to 98%, consuming approximately 41 minutes and returning AUC equal to 0.81.

The tests were performed with 10, 30, 50 and 100 epochs and the number of images used in the training were 50, 80, 100, 150, 200, 300 and 500 for both genders. A difference can be seen here between the accuracy results and the AUC score between the genders. This difference in accuracy, seen to be greater for males, is understandably due to the fact that there is one less class for classification by the CNN when compared to the female classes. The AUC has higher scores for females seen as there are more classes to compose the curve when compared to the male classes. Another important data that the CNN returns is the accuracy for the validation set at the time of the CNN training.

During its training, the CNN performs validations to adjust the weights and returns the accuracy value for this validation set. For males, the validation had an accuracy of 98%, and for females it was 89%. Better numbers for accuracy and AUC

Table 4. Confusion matrix of the CNN training for females

Actual value	Predicted value					Accuracy	F1 Score
	18-35	36-50	51-60	61-70	+71		
18-35	80	1	9	0	0	89%	0.90
36-50	3	265	1	5	0	97%	0.94
51-60	5	18	473	12	0	93%	0.91
61-70	0	6	38	351	0	89%	0.91
+71	0	2	13	7	100	82%	0.90

Table 5. Confusion matrix of the CNN training for males

Actual value	Predicted value				Accuracy	F1 Score
	18-35	36-50	51-70	+71		
18-35	111	1	0	0	99%	0.99
36-50	1	187	7	1	95%	0.97
51-70	1	2	260	2	98%	0.98
+71	0	0	0	60	100%	0.98

were obtained with training configurations using more epochs and more images, but the time consumed was much higher. For both genders, the time taken to reach these values was 5 times greater, however the accuracy and AUC scores were about 5% higher for females and between 1% and 2% for males. With the CNN trained, the system saves the model and weights, separated by gender, to be used in the classification.

3.4.2 Classification Results and the CNN Training

For female patients, the trained CNN, with the training database, returned an accuracy equal to 94%, using a total of 1389 images, distributed as follows: class 18-35 = 90 images; 36-50 = 274 images; 51-60 = 508 images; 61-70 = 395 images; +71 = 122 images. The CNN training resulted in the confusion matrix in Table 4.

For male patients, also with the training database, the CNN returned an accuracy equal to 98% with a database containing 633 images, distributed as follows: class 18-35 = 112 images; 36-50 = 196 images; 51-70 = 265 images; +71 = 60 images. As shown in the confusion matrix in Table 5.

As we can observe in the different classes of tables 4 and 5, the value of the F1 Score was above 0.9. We also calculated the F1 weighted average resulting 0.91 and 0.97 for that tables. These results show a good and stable classifier.

After the CNN was trained, a manual test was performed considering all the images in the database of test from both genders, classifying image by image. The result is shown by means of a confusion matrix, Table 6 and Table 7, for females and males respectively. The F1 Score values presente in these tables were considered good, despite observed mistakes, as mentioned below. The F1 weighted average was also calculated, resulting 0.90 and 0.86, showing a good and stable classifier, evaluated with unknow images.

The results showed some interesting data. In the manual classification for the class 36-50 of the male gender, three exams showed all images classified by the CNN in discrepant

Table 6. Confusion matrix of the manual test for females

Actual value	Predicted value					Accuracy	F1 Score
	18-35	36-50	51-60	61-70	+71		
18-35	76	0	14	0	0	84.4%	0.92
36-50	0	254	19	1	0	92.7%	0.93
51-60	0	7	474	27	0	93.3%	089
61-70	0	7	33	353	2	89.3%	0.91
+71	0	2	19	3	98	80.3%	0.88

Table 7. Confusion matrix of the manual test for males

Actual value	Predicted value				Accuracy	F1 Score
	18-35	36-50	51-70	+71		
18-35	112	0	0	0	100%	0.90
36-50	11	132	45	8	67.3%	0.77
51-70	15	16	240	10	85.4%	0.85
+71	0	0	0	120	100%	0.93

classes of the age of the patient. As the data were anonymized during collection, it was not possible to find the medical records of these patients to verify if they had any associated risk factors or if they were affected by osteoporosis. Thus, the low accuracy of this range is due to these three anomalous cases. In the classes of extremities, 18-35 and +71, 100% of accuracy was obtained; this may be linked to the reduced number of images obtained. For females, the extremity classes, 18-35 and +71, had fewer images, but here the results were different in terms of accuracy: 84.4% and 80.3% respectively. Classes 36-50 and 51-60 had a very close precision to the training accuracy returned by the CNN. When observing the menopausal period (51-60 years), the CNN was able to generalize and classify this age range of the samples well, revealing the highest percentages of accuracy. The results show good accuracy during training; however, when considering the average of at least 5 sample images per exam in the analysis, the results obtained are much more robust.

It is a matter of calculating the average of the percentages of each class of the selected slices, which, in a way, reduces the importance of a disparate assessment in the universe of the analyzed layers. This minimizes adverse results due to samples which are low quality, close to implants, or that present some structural anomaly of the bone. Given the above-mentioned, a new test was carried out: choosing 10 random exams, giving preference, when possible, to the validation base; in each exam, 5 coronal slices were analyzed. This test was carried out for both sexes and aimed to show the assertiveness of classification by the CNN; Table 8 shows data for females, and Table 9 for males.

It can be observed that there was an accuracy of 100% for females and only one error for males – which is one of the 3 anomalous exams mentioned above.
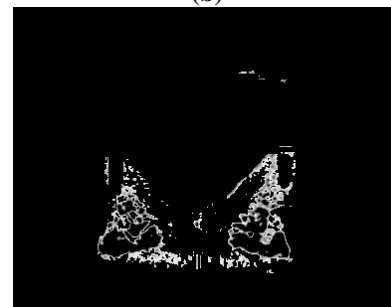
We understand that the proposed model was able to achieve generalization capacity, analyzing the texture of the image and not its shape, which is a strong indication that CNN has
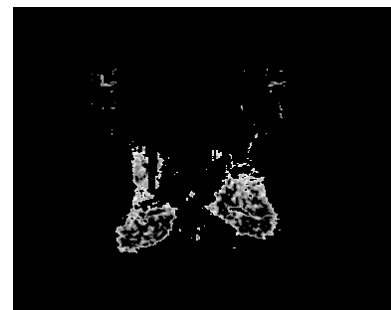
(a)
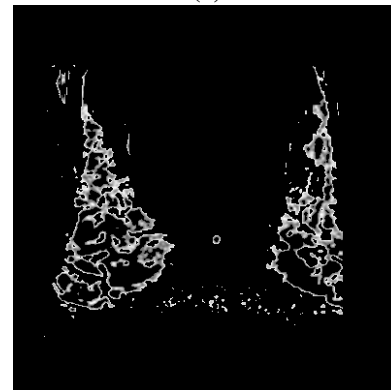
(b)

(c)

Figure 4. Examples of already segmented samples with varied formats of the ROI

learned to classify patients by their age through the texture of the trabecular bone structure. In addition, from the training results, it is clear that the evaluation indices, especially accuracy, demonstrate that overfitting does not happen. Figure 4 shows that the CNN does not consider the ROI format of the segmented image as a preponderant factor in image classification. In this Figure there are 3 images (a), (b) and (c) with different shapes, dimensions, and sizes, but belonging to the same class – in this case, 51-70 males – and the CNN classified them correctly in the performed test. The values shown in this study should not be used for the diagnosis of osteopenia or osteoporosis when presenting images from CBCT scans classified in classes distinct from the age of the patient. However, this result can be considered as it points out some anomaly in the bone structure of the patient and thus guide them to seek medical help specialized in bone density assessment.

Table 8. Validation of exams for females

Classes	Exam Quantity	Hits	Errors
18-35	6	6	0
36-50	10	10	0
51-60	10	10	0
61-70	10	10	0
+71	9	9	0

Table 9. Validation of exams for males

Classes	Exam Quantity	Hits	Errors
18-35	10	10	0
36-50	10	9	1
51-70	10	10	0
+71	4	4	0

3.4.3 Data Presentation to the User

The software presents as results the percentage values of belonging, associated with the age classes of each sample and the average of these values for all images sampled in a CT, in addition to statistical data. Figure 5 shows the graphical interface that presents these results.

4. Conclusion

Bone density is directly linked to quality of life, seen as the lower the bone density, the greater the probability of suffering bone fractures. There is evidence of a correlation between the density of the mandibular bone and the density of the femoral neck, a region used in DXA scans for the diagnosis of osteoporosis [28]. As a DXA is not accessible to a large part of the Brazilian population due to its high cost and the fact that the equipment available in the Brazilian public health system (SUS) is poorly distributed throughout the national territory, making its accessibility more difficult, it is necessary to use other sources of information to assess the bone structures of patients.

The mandibular bone is an object of work for dentists and CBCT scans are often used in dental implant treatments. In this study, a new methodology was proposed to estimate the relationship between the age of patients and the age of the bone structure of their mandibles through the analysis of CBCT images by a CNN. The CNN, in turn, classifies the trabecular structure of the mandible into one of the age groups, as shown in Tables 1 and 2.

A software named Jaw System Age Group X was built, which incorporates the proposed methodology and classifies the age of the trabecular bone of the mandible using images from CBCT scans in different age groups according to the gender of the patient. The CNN, which is incorporated into the software, proved to be robust and to have a good generalization ability after its training. In the CNN training, the accuracy returned was equal to 94%, and the AUC value was equal to 0.94 for females; for males, the accuracy was 98% and the AUC was 0.81. These results are obtained through the individual analysis of the slices. The average result cal-

Figure 5. Graphical interface for presentation of results

culated from the analysis of at least five consecutive slices of the mandible of a patient was excellent, providing accuracy values close to 100%, for both genders, as shown in Table 8 and Table 9.

As future work, it is proposed include machine learning interpretability experiments to see how and why the model is making its predictions.

In conclusion, the software developed is a low-cost alternative with the potential to assist healthcare professionals in analyzing the bone density of patients, being able to serve as a filtering criterion in performing diagnostic tests, such as the DXA.

5. Author contributions

Matheus Raffael Simon: contributed to the development of this study; acted in the construction of the concept and methods for the conception of the research; participated in data acquisition and analysis; wrote, read and approved the final manuscript.

Adair Santa Catarina: contributed to the development of this study; assisted in the construction of the concept and methods for designing the research; participated in data acquisition and analysis; wrote, read and approved the final manuscript.

Adriane Yaeko Togashi: contributed to the development of this study; suggested the concept and methods for designing the research; participated in acquiring and analyzing the data; read and approved the final manuscript.

References

[1] ZERBINI, C. Osteoporose — Entrevista. *Drauzio Varella*, 2012. Disponível em: ⟨https://drauziovarella.uol.com.br/entrevistas-2/osteoporose-entrevista/⟩. Acesso em: 23 jul. 2020.

[2] GUIMARÃES, M. F. B. R. et al. Qual o melhor ponto de corte de índice de massa corporal para diagnosticar a obesidade em mulheres com artrite reumatoide? um estudo que usa a composição corporal pela absorciometria com raios x de dupla energia. *Revista Brasileira de Reumatologia*, São Paulo, v. 57, n. 4, p. 279–285, jul. 2017.

[3] YAZBEK, M. A.; NETO, J. F. M. Osteoporose e outras doenças osteometabólicas no idoso. *Einstein*, São Paulo, v. 6, n. 1, p. 74–78, mar. 2008.

[4] COOPER, C.; III, L. J. M. Epidemiology of osteoporosis. *Trends in Endocrinology & Metabolism*, Nova Iorque, v. 3, n. 6, p. 224–229, ago. 1992.

[5] VESTERBY, A. Star volume of marrow space and trabeculae in iliac crest: sampling procedure and correlation to star volume of first lumbar vertebra. *Bone*, Nova Iorque, v. 11, n. 3, p. 149–155, jan. 1990.

[6] GIORDANO, V. et al. Alterações na estrutura óssea relacionadas à idade. *Revista do Colégio Brasileiro de Cirurgiões*, Rio de Janeiro, v. 43, n. 4, p. 276–285, jul. 2016.

[7] RADOMINSKI, S. et al. Osteoporose em mulheres na pós-menopausa. *Revista Brasileira de Reumatologia*, São Paulo, v. 44, n. 6, p. 426–434, dez. 2004.

[8] WOWERN, N. V.; KLAUSEN, B.; KOLLERUP, G. Osteoporosis: a risk factor in periodontal disease. *Journal of periodontology*, Chicago, v. 65, n. 12, p. 1134–1138, dez. 1994.

[9] SANTOS, M. K. et al. Artificial intelligence, machine learning, computer-aided diagnosis, and radiomics: advances in imaging towards to precision medicine. *Radiologia brasileira*, São Paulo, v. 52, n. 6, p. 387–396, nov. 2019.

[10] DANGETI, P. *Statistics for machine learning*. 1. ed. Reino Unido: Packt Publishing Ltd, 2017.

[11] JUDAS, F. et al. *Estrutura e dinâmica do tecido ósseo*. Portugal: Clínica Universitária de Ortopedia, HUC, 2012.

[12] RODRIGUES, R. D. et al. Manejo cirúrgico de fratura de mandíbula: relato de caso. *Revista Da Faculdade de Odontologia-UPF*, Passo Fundo, Brasil, v. 23, n. 3, p. 343–347, set. 2018.

[13] KUMAR, V. et al. *Robbins patologia básica*. 9. ed. Rio de Janeiro: Elsevier, 2013. v. 10.

[14] RENA, R. J. *A mulher e a osteoporose: como prevenir e controlar*. 5. ed. São Paulo: Estante de Saúde, 2019.

[15] DORIA, G. Osteoporose: Saiba quais fatores influenciam a doença: segundo a organização mundial da saúde, 10 milhões de brasileiros sofrem com o problema. *Pleno News*, 2017. Disponível em: ⟨https://pleno.news/brasil/osteoporose-saiba-quais-fatores-influenciam-a-doenca.html⟩. Acesso em: 01 jan. 2021.

[16] VIEIRA, R. A. A atuação da fisioterapia na prevenção de quedas em pacientes com osteoporose senil. *Fisioterapia Brasil*, São Paulo, v. 3, n. 2, p. 72–78, mar. 2002.

[17] LUIZE, D. S. et al. A influência da osteoporose na implantodontia. *Arquivos em Odontologia*, Belo Horizonte, v. 41, n. 2, p. 92–105, abr. 2005.

[18] GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. *Deep learning*. Cambridge: MIT press, 2016.

[19] LECUN, Y.; BENGIO, Y.; HINTON, G. Deep learning. *Nature*, Londres, v. 521, n. 7553, p. 436–444, maio 2015.

[20] PENHA, D. d. P. Rede neural convolucional aplicada à identificação de equipamentos residenciais para sistemas de monitoramento não-intrusivo de carga. Dissertação (Mestrado em Engenharia Elétrica) - Programa de Pós-graduação em Engenharia Elétrica (PPGEE), Instituto de Tecnologia, Universidade Federal do Pará, Belém, p. 71, 2018.

[21] TOGASHI, A. et al. Bone microarchitecture by dentistry digital x-ray (bμa-ddx) software: A pilot study of the analysis of bone density using digital dental x-rays. *Current Research*

in Dentistry, Emirados Árabes Unidos, v. 6, n. 2, p. 18–26, fev. 2015.

[22] LEE, J.-S. et al. Osteoporosis detection in panoramic radiographs using a deep convolutional neural network-based computer-assisted diagnosis system: a preliminary study. *Dentomaxillofacial Radiology*, Reino Unido, v. 48, n. 1, p. 1–8, jan. 2019.

[23] TANG, C. et al. Cnn-based automatic detection of bone conditions via diagnostic ct images for osteoporosis screening. *Arxiv preprint arXiv: 1910.06777*, [S. l.], p. 1–26, out. 2019.

[24] BOTTEGA, P. Correlação dos fatores locais e sistêmicos com a densidade óssea dos maxilares obtido através de tomografia computadorizada cone-beam. 2021. Dissertação (Mestrado em Odontologia) - Programa de Pós-Graduação em Odontologia, Centro de Ciências Biológicas e da Saúde, Universidade Estadual do Oeste do Paraná, Cascavel, p. 39, 2021.

[25] BORGES, M. S.; MUCHA, J. N. Avaliação da densidade óssea para instalação de mini-implantes. *Dental Press Journal of Orthodontics*, Brasil, v. 15, n. 6, p. e1–e9, dez. 2010.

[26] HIASA, K. et al. Preoperative computed tomography-derived bone densities in hounsfield units at implant sites acquired primary stability. *International Scholarly Research Notices*, EUA, v. 2011, p. 1–5, nov. 2011.

[27] MARIAUX, S. et al. Is preoperative glenoid bone mineral density associated with aseptic glenoid implant loosening in anatomic total shoulder arthroplasty? *BMC musculoskeletal disorders*, Reino Unido, v. 22, n. 1, p. 1–9, jan. 2021.

[28] SCHEIBEL, P. C.; RAMOS, A. L.; IWAKI, L. C. V. Is there correlation between alveolar and systemic bone density? *Dental press journal of orthodontics*, Brasil, v. 18, n. 5, p. 78–83, out. 2013.