

Impact of Feature Selection on Clustering Images of Vertebral Compression Fractures

Raquel M. Candido¹, Rafael S. Del Lama¹, Natália S. Chiari²,
Marcello H. Nogueira-Barbosa², Paulo M. de Azevedo Marques², Renato Tinós¹

¹Department of Computing and Mathematics
FFCLRP, University of São Paulo (USP) – Ribeirão Preto, SP – Brazil

²Department of Medical Imaging, Hematology and Clinical Oncology
FMRP, University of São Paulo (USP) – Ribeirão Preto, SP – Brazil

{raquel.candido, rafael.lama, natalia.chiari}@usp.br

marcello@fmrp.usp.br, pmarques@fmrp.usp.br, rtinos@ffclrp.usp.br

Abstract. *Non-traumatic Vertebral Compression Fractures (VCFs) are generally caused by osteoporosis (benign VCFs) or metastatic cancer (malignant VCFs) and the success of the medical treatment strongly depends on a fast and correct classification of VCFs. Recently, methods for computer-aided diagnosis (CAD) based on machine learning have been proposed for classifying VCFs. In this work, we investigate the problem of clustering images of VCFs and the impact of feature selection by genetic algorithms, comparing the clustering i) with all features and ii) with feature selection through the purity results. The analysis of the clusters helps to understand the results of classifiers and difficulties of differentiating images of different classes by an expert. The results indicate that features selection improved the separability of clusters and purity. Feature selection also helps to understand which attributes are most important for analysing the images of vertebral bodies.*

1. Introduction

The spine is the central axis of the human body. Even though the spine is flexible, injuries affecting one or more of the 33 vertebrae can have serious effects on the life quality of patients [Vasconcelos 2004]. Non-traumatic Vertebral Compression Fractures (VCFs) generally result from osteoporosis (benign VCFs) and metastatic cancer (malignant VCFs). The diagnosis of non-traumatic VCFs is made by radiologists generally through the use of magnetic resonance images (MRI), allowing the professional to differentiate malignant and benign VCFs [Tehranzadeh and Tao 2004].

There is an increasing interest in the use of Machine Learning techniques in medical image analysis. Several reasons can be cited for this interest, such as the small number of experienced professionals, compared to the large number of medical images generated every day, and the possible occurrence of mistakes caused by stress and fatigue [Azevedo-Marques 2001], [Shen et al. 2017]. Concerning the analysis of images of VCFs, early works proposed methods for classifying the shape of VCFs [Genant et al. 1993] and diagnosing osteoporosis [Kasai et al. 2006] in images. In [Ribeiro et al. 2012], Gabor filters and artificial neural networks were applied for segmentation of vertebral bodies, that were then classified using the semiquantitative grading method proposed in [Genant et al. 1993].

The use of machine learning was proposed in [Frighetto-Pereira et al. 2015] for classifying benign and malignant VCFs. Several texture, gray-level, and shape features extracted from MRI of segmented vertebral bodies were used for training and testing different classifiers. Feature selection was applied by using a *wrapper* method present in WEKA [Hall et al. 2009]. Despite using a simple method for feature selection (little information about the method and its use is informed by the authors), the work indicated that the use of feature selection increased the classification performance.

In [Raineri 2018], two feature selection methods were compared using the same dataset employed in [Frighetto-Pereira et al. 2015], but considering 3 classes (normal, benign, and malignant) instead of just two (benign and malignant). The methods were based on genetic algorithms and sequential forward selection, and were also compared to using principal component analysis to combine features. K-nearest neighbours, multilayer perceptron, and support-vector machines were used for classifying the MRI. The same dataset was also used in [Lama 2018], but using a different approach: a convolutional neural network was used to classify the raw images, i.e., features were not extracted *a priori*. Despite some good results, a limitation of the deep learning approach is the necessity of having a large dataset for training the convolutional neural network.

In this work, we investigate the problem of clustering images of VCFs. According to our knowledge, our research work is the first investigating the use of clustering algorithms in the image analysis of VCFs. In fact, the use of clustering for analysing MRI in similar medical images analysis problems seems to be new. However, clustering is important to understand better the difficulties of applying machine learning in computer-aided diagnosis. The decision regions formed by a classifier are, many times, strongly related to clusters of data. Understanding how the MRIs of different classes are grouped in clusters can help designing new classifiers and features for automated medical analysis. Clustering can also help understanding the difficulties of specialists in differentiating images from different classes.

Here, the impact of feature selection on clustering images of VCFs is investigated using k-means, the most popular clustering method used on scientific applications [Berkhin 2002]. As some features are irrelevant and redundant for classification and clustering, feature selection is important and should help understanding the importance of different attributes for clustering the vertebral bodies images. A genetic algorithm is used for feature selection; evolutionary algorithms have been successfully used in many feature selection problems [Xue et al. 2015, Hruschka et al. 2009].

This paper is organized as follows. The dataset, composed of segmented images of vertebral bodies obtained from MRI, is described in Section 2. The methodology used for clustering and feature selection is described in Section 3. Section 4 shows the experimental results, while Section 5 concludes the paper.

2. Dataset

A set of MRI of the spine of 61 patients at HCFMRP/USP (“*Hospital das Clínicas da Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo*”) was used here. These patients were diagnosed with VCF by radiologists of HCFMRP/USP. The images are from the median sagittal plane (central section) weighted in the T1 contrast sequence. They were acquired following the clinical routine protocol in a Philips Achieva 1.5T

MRI system. The vertebral bodies were manually segmented from the spine images. A radiologist labeled each one of the vertebral bodies. The spine images for the patients with VCFs also contained some vertebral bodies labelled as normal. The examples for the normal class, used here, were obtained from patients with benign VCFs. A total of 189 examples were selected after the segmentation process (see Table 1).

Table 1. Labels for the examples in the dataset.

Class	Amount
Benign VCF	54
Malignant VCF	46
Normal	89
Total	189

Radiomic features were extracted from the segmented original 3-D MRI using the *PyRadiomics* software [Van Griethuysen et al. 2017]. A total of 106 features were extracted with the following distribution:

- First Order Statistics: 19 attributes
- Shape-based (3-D): 14 attributes
- Gray Level Co-occurrence Matrix: 23 attributes
- Gray Level Size Zone Matrix: 16 attributes
- Gray Level Run Length Matrix: 16 attributes
- Neighboring Gray Tone Difference Matrix: 4 attributes
- Gray Level Dependence Matrix: 14 attributes

The set of segmented images, with some modifications, was originally used in [Frighetto-Pereira et al. 2015]. The use of this set of images for research purposes was approved by The Research Ethics Committee of HCFMRP/USP. Here, the features were normalized by using the median and the standard deviation. The original dataset used here contains 189 examples with 106 attributes.

3. Methodology

In order to investigate the impact of feature selection in clustering, two sets of experiments were performed:

- clustering with the original dataset, i.e., considering examples with all 106 attributes;
- clustering with datasets where attributes were obtained by feature selection.

Figure 1 shows the the framework for the analysis of the impact of feature selection in clustering. The next two sections show the methods for feature selection and clustering.

3.1. Feature selection

Genetic algorithms (GAs) have been successfully applied to the feature selection problem [Xue et al. 2015, Hruschka et al. 2009]. Despite the longer time required when compared to greedy strategies, GAs find generally better local optima when applied to feature selection. We use here the wrapper approach, i.e., the GA searches for good subsets of attributes for a specific classifier. Given a classifier and a dataset, a chromosome of the

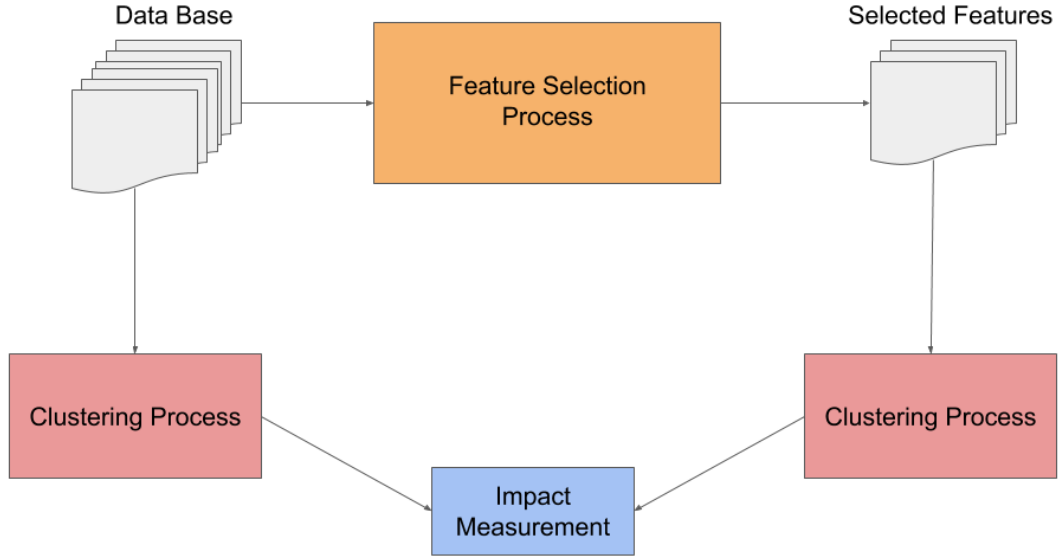


Figure 1. Framework for the analysis of the impact of feature selection in clustering.

GA, $\mathbf{x} \in \mathbb{B}^N$, indicates a subset of selected features (attributes). The total number of features in the dataset is indicated by N ; here, the original dataset has $N=106$ attributes. In the chromosome, $x_i = 1$ indicates the presence of the i -th attribute, while $x_i = 0$ indicates that it is not present in the selected features subset. Bit flip mutation, with rate p_m , and two-point crossover, with rate p_c , are applied. Tournament selection and elitism are used for the selection of individuals.

The initial population is randomly generated, where each individual contains a random number of ones (number of features) between $0.25N$ and $0.75N$. The fitness function is given by:

$$f(\mathbf{x}) = acc(\mathbf{x}) + \left(1 - \frac{\sum_{i=1}^N x_i}{N}\right) \quad (1)$$

where $acc(\mathbf{x})$ is the accuracy of the classifier when attributes given in \mathbf{x} are used. The classifier is a Multilayer Perceptron (MLP). Using the MLP is important because one of our objectives, when analysing the clusters formed when feature selection is applied, is to understand better the decision regions formed by this kind of artificial neural network. In Eq. 1, $acc(\mathbf{x})$ is computed considering *5-folds* cross validation. In other words, each time an individual is evaluated, the MLP is trained and tested 5 times, and $acc(\mathbf{x})$ is computed over all test sets. The MLP has a fixed architecture and is trained by using Backpropagation. The second term in Eq. 1 is higher for subsets with fewer attributes. Thus, the GA should find subsets with few attributes and that result in classifiers with higher accuracy.

When running the GA, different best solutions are found for different runs. After running the GA for all runs, we compute the frequency of the selected features (attributes) in the best solutions. One best solution is obtained for each run. The clustering analysis is made for attributes selected by the GA for:

- attributes with frequency greater than 50%
- attributes with frequency greater than 20%
- attributes with frequency greater than 10%

Not all vertebral bodies were used in evaluating the individuals of the GA. We reserved 20 vertebral bodies [Raineri 2018] to compose the dataset to perform the clusterings after feature selection.

3.2. Clustering

We consider hard partition clustering, i.e., each example (image here) belongs to only one cluster [Xu 2005]. The technique used for clustering the images here is k-means [MacQueen 1967] given the large usability due the practicality and simplicity of the technique. In k-means, the sum of square Euclidean distances from the objects to the centroids of the clusters is minimized. This is done by the successive computation of the centroids of k clusters and reallocation of examples to the closest centroids. The number of clusters, k , is fixed in k-means. Another important property of k-means is that it generates hyper-spherical clusters.

In order to analysis the clusters, the purity is computed. This is possible because the classes for the problem investigated here are known. When defining k , the largest purity for running k-means with $2 \leq k \leq \sqrt{T}$ is selected, where $T = 189$ is the number of examples in the dataset. The purity for cluster C is given by:

$$Purity(\Omega, \mathbf{C}) = \frac{1}{T} \sum_{k=2}^T \max_j |\omega_k \cap \mathbf{c}_j| \quad (2)$$

where $\Omega = \omega_1, \omega_2, \dots, \omega_k$ indicates the set of clusters and $\mathbf{C} = \mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_j$ the set of labels (class). For each cluster, the class with more elements in the cluster is used for computing the purity.

Here, the purity is also used for analysing the impact of feature selection. In this way, the different partitions obtained by running k-means for the original dataset (i.e, without feature selection) and for the datasets obtained by feature selection considering different runs of the GA are analysed. We also present 2-D plots for the different partitions obtained by running k-means. Principal component analysis (PCA) is employed to represent the distribution of the examples in 2 dimensions, i.e., 2 principal components are used to plot the vertebral bodies on the bidimensional space. Also, the proportion of objects for different classes in each cluster was observed in the experiments.

4. Results

The parameters of the GA were: population size equals to 100 individuals, number of generations equals to 500, $p_c = 0.6$, and $p_m = \frac{1}{N}$. The number of runs is 10. The parameters of the MLP were: one hidden layer with 10 neurons, hyperbolic tangent activation function, learning rate equals to 0.3, and 200 training epochs.

4.1. Experimental Results

For feature selection, the average number of selected attributes by the GA was 5 and the average accuracy for the best solutions was 0.82. The best solution (considering all 10

runs) selected 4 attributes, with average accuracy equals to 0.849. Table 2 shows the frequency of the most frequent attributes. The two most selected attributes are shape features (Table 2).

We show now the results for finding the best value of k for k-means. We consider datasets with the following number of attributes selected by the GAs: 9 (frequency greater than 10%), 4 (frequency greater than 20%), and 3 (frequency greater than 50%). Table 3 shows the purity for the different partitions when running k-means with different k . In general, the best results were obtained for large values of k , what is a expected result. In the next, we consider the configuration that resulted in the highest purity value for $2 \leq k \leq \sqrt{189}$; the set of 4 attributes and $k = 13$ that obtained 0.76 of purity.

Figure 2 shows the 2-D plots for $k = 13$ for the dataset with all attributes and for the dataset with the selected features. In the following, the results for the datasets with all attributes and with the selected features are show.

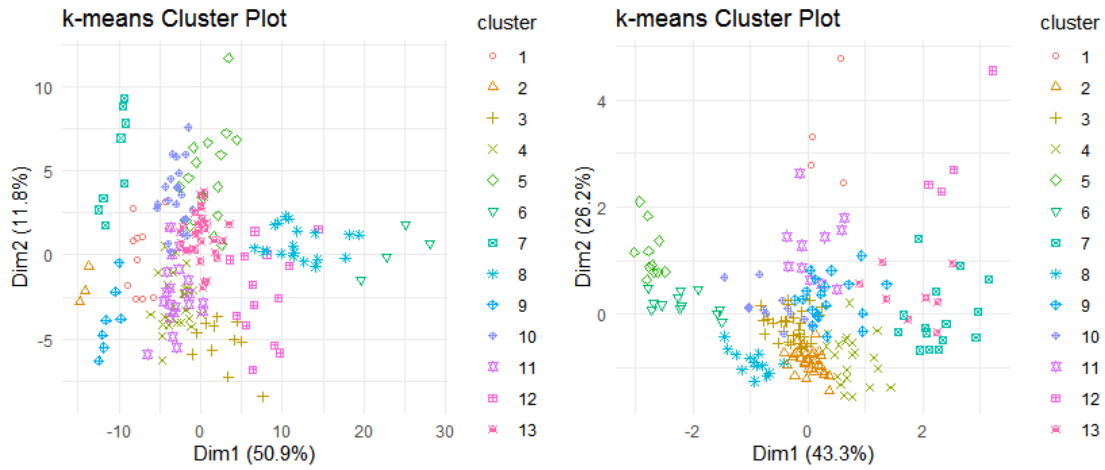


Figure 2. Cluster plots for the partitions obtained by k-means, considering datasets with a) all attributes and b) selected attributes for $k = 13$.

4.1.1. All attributes, $k = 13$

Table 4 shows the majority class for the examples inside each cluster (group) for $k = 13$ in the experiment with all attributes. Most of the examples classified as benign VCF, malignant VCF and normal belong to clusters 5, 11 and 13, respectively (these results are not shown here). Malignant bodies are not found in clusters 6 and 8, although clusters 2 and 9 only have malignant examples. Table 4 also shows the purity and the majority class for each group. Clusters 1 and 6 do not contain the majority class well defined because the two class have the same amount of vertebral bodies.

4.1.2. Selected features, $k = 13$

Table 4 shows the majority class for $k = 13$ in the experiment with selected features (4 attributes). Most of the examples classified as benign VCF, malignant VCF and non fractured are in clusters 10, 7 and 3, respectively (these results are not shown here). Malignant

Table 2. Frequencies for the attributes selected in more than 10% of the runs of the GA.

Feature	Frequency
<i>glcm_ClusterTendency</i>	20%
<i>glcm_Imc1</i>	20%
<i>shape_MajorAxis</i>	100%
<i>shape_SurfaceVolumeRatio</i>	90%
<i>glcm_Imc2</i>	30%
<i>firstorder_Median</i>	60%
<i>gldm_GrayLevelVariance</i>	20%
<i>firstorder_StandardDeviation</i>	20%
<i>glcm_Correlation</i>	20%

bodies are not found in clusters 1, 3, 5 and 6; however, cluster 7 contains only malignant examples. Table 4 also shows the purity and the majority class for each group. Cluster 1 do not contain the majority class well defined because the benign VCF and non fractured classes have the same amount of vertebral bodies (50%).

Table 3. Purity for partitions obtained by k-means with different k . Results for all attributes and with feature selection (for 9, 4 and 3 selected attributes) are shown. The best values are in bold.

k	All attributes	9 attributes	4 attributes	3 attributes
2	0.47	0.47	0.47	0.47
3	0.57	0.52	0.58	0.51
4	0.58	0.52	0.60	0.61
5	0.57	0.54	0.61	0.62
6	0.60	0.57	0.60	0.62
7	0.60	0.58	0.60	0.69
8	0.60	0.60	0.60	0.70
9	0.60	0.63	0.60	0.69
10	0.65	0.63	0.70	0.72
11	0.61	0.64	0.70	0.74
12	0.62	0.65	0.71	0.75
13	0.67	0.69	0.76	0.75
Mean	0.60	0.59	0.63	0.66

4.2. Analysis of the Clustering Results

We analyse the clusters results regarding some of the classification results obtained by KNN and reported in [Raineri 2018]. The results are presented regarding the test set, that contains 20 images of vertebral bodies for patients not seen during training by KNN. The same dataset used by [Raineri 2018] is used here for clustering; however, it is important to highlight that the images of the test set were not used on the feature selection process. The results for the clustering experiments are also analysed regarding another level of classification for the malignant VCFs. In this level, the causes for the malignant VCFs

Table 4. Results for the experiments with all attributes and $k = 13$. *Proportion* is for examples of the majority class in the cluster and *Total* indicates the number of vertebral bodies (examples) in the cluster.

Group	Majority Class	Proportion	Purity	Total
1	Benign/Malignant* (4)	36%	0.021	11
2	Malignant (3)	100%	0.016	3
3	Normal (7)	64%	0.037	11
4	Normal (16)	67%	0.085	24
5	Benign (10)	63%	0.053	16
6	Benign/Normal* (2)	50%	0.011	4
7	Malignant (7)	88%	0.037	8
8	Normal (16)	84%	0.085	19
9	Malignant (7)	100%	0.037	7
10	Normal (14)	67%	0.074	21
11	Malignant (11)	29%	0.058	21
12	Benign (9)	64%	0.048	14
13	Normal (21)	70%	0.111	30

*groups with the same amount of vertebral bodies.

(e.g., consequence of breast cancer) are analysed. The labels for the classification of the causes of VCFs were informed by specialists.

In [Raineri 2018], KNN incorrectly classified 3 vertebral bodies of the test set. In the following, each one of them are analysed regarding the clustering results. For the identification of the examples, P is used for identifying the patient and L for the lumbar vertebral body. The incorrectly classified examples were:

- P20L1 was classified as malignant, but the real class is benign. In the clustering results:
 - for all attributes: P20L1 belongs to group 10 and the majority class in this cluster is not well defined: the cluster has the same amount of benign VCF and malignant VCF. Group 10 has bodies classified as multiple myeloma (1) and chronic leukemia (2).
 - for selected features: P20L1 belongs to group 4, with malignant majority class. Group 4 has malignant bodies classified as multiple myeloma (6), breast carcinoma (2), others (2), oropharyngeal carcinoma (1), cholangiocarcinoma (1) and paraganglioma (1).

The cluster for the experiment with selected features is mainly composed by examples in the malignant class. This helps explaining why the example was incorrectly classified as malignant. Besides, the clusters contains malignant examples of the multiple myeloma type. In general, myeloma is not easily identified in the images with VCFs.

- P33L1 was classified as normal, but the real class is benign. In the clustering results:
 - for all attributes: P33L1 belongs to group 13, where the majority class is normal.
 - for selected features: P33L1 belongs to group 11, where the majority class is benign.

Table 5. Results for the experiments with selected attributes and $k = 13$.

Group	Majority Class	Proportion	Purity	Total
1	Benign/Normal* (2)	50%	0.011	4
2	Normal (21)	70%	0.111	30
3	Normal (22)	76%	0.116	29
4	Malignant (13)	62%	0.069	21
5	Normal (7)	64%	0.037	11
6	Normal (11)	92%	0.058	12
7	Malignant (15)	100%	0.079	15
8	Normal (9)	69%	0.048	13
9	Normal (15)	79%	0.079	19
10	Benign (11)	92%	0.058	12
11	Benign (9)	82%	0.048	11
12	Malignant (3)	75%	0.016	4
13	Benign (5)	63%	0.026	8

*groups with the same amount of vertebral bodies.

In the experiment for all attributes, P33L1 belongs to a cluster where the majority class is normal, what helps explaining why this example was incorrectly classified as normal. However, for selected features, P33L1 belongs to a cluster where the majority class is benign (true class), what is reasonable, considering a selective set of features.

- P24L1 was classified as normal but the real class is malignant. This example is of breast carcinoma type. In the clustering results:
 - for all attributes: P24L1 belong to group 5, where the majority class is benign. Group 5 has malignant examples of multiple myeloma type (2);
 - for selected features: P24L1 belong to group 11, where the majority class is benign. Group 11 has malignant examples of types multiple myeloma (1) and breast carcinoma (1). One can observe that group 11 is the same group that contains P33L1.

The results indicated that P24L1 is in clusters where the majority class is benign. However, the resulting classification is normal, despite of the example being of malignant class. One can also observe that multiple myeloma bodies were in the same cluster of multiple myeloma and breast carcinoma.

5. Conclusion

The cluster analysis indicates that features selection improved the external evaluation (Table 3), indicating that the separability of the clusters was improved. The best results for purity was for the same value of k (13). For the experiments with all attributes, the purity was 0.67, while it was 0.76 for the experiments with selected features. Feature selection also helps to understand which attributes are most important when specialists analyze the vertebral bodies images. The two attributes that were most selected by the GA were shape features (Table 2). When images of vertebral bodies are analysed by radiologists, shape features are important in class differentiation. This happens because the shape of non-fractured bodies is well-defined, while benign and malign VCFs cause alterations in the shape [Raineri 2018].

The analysis of the clusters helps to understand the results of classifiers and difficulties of differentiating images of different classes by an expert. The analysis of the clusters helped to understand the classifier errors in the test set. Three vertebral bodies (examples) were incorrectly classified by KNN in the experiments presented in [Raineri 2018]. In the experiments with selected features, two of the examples were classified as the class that belongs to the majority class of the respective cluster. Feature selection allowed to obtain better results when the misclassified examples were qualitatively analysed. The qualitative analysis also allowed to identify the types of malignant VCFs in the clusters, allowing to help understanding the similarities in the images. Some images will be analyzed by a radiologist in the future in order help understanding better the difficulties of differentiating the classes.

The setting of the parameters used in feature selection and the analysis of classification results through the clustering outcomes was based on previous experiments [Raineri 2018]. In the experiments, the impact of feature selection was positive. Regarding [Raineri 2018], some different features were selected. This can be explained by the use of z-score instead min-max normalization[Alelyani et al. 2018] during preprocessing.

Here, we used k-means for clustering the images of vertebral bodies. Algorithm k-means creates hyper-spherical clusters. For future works, we should investigate a density-based clustering method [Tinós et al. 2018]. This method can create clusters with arbitrary shapes, and can automatically optimize the number of clusters.

Acknowledgements

This work was partially supported by São Paulo Research Foundation (FAPESP), under grants 2013/07375-0 and 2019/07665-4, and National Council for Scientific and Technological Development (CNPq), under grant 305755/2018-8.

References

- Alelyani, S., Tang, J., and Liu, H. (2018). *Feature Selection for Clustering: A Review*, pages 29–60.
- Azevedo-Marques, P. M. (2001). Diagnóstico auxiliado por computador na radiologia. *Radiologia Brasileira*, 34(5):285–293.
- Berkhin, P. (2002). *Survey of Clustering Data Mining Techniques*. Accrue Software.
- Frighetto-Pereira, L., Menezes-Reis, R., Metzner, G. A., Rangayyan, R. M., Nogueira-Barbosa, M. H., and Azevedo-Marques, P. M. (2015). Classification of vertebral compression fractures in magnetic resonance images using shape analysis. In *2015 E-Health and Bioengineering Conference (EHB)*, pages 1–4. IEEE.
- Genant, H. K., Wu, C. Y., Van Kuijk, C., and Nevitt, M. C. (1993). Vertebral fracture assessment using a semiquantitative technique. *Journal of bone and mineral research*, 8(9):1137–1148.
- Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., and Witten, I. H. (2009). The weka data mining software: an update. *ACM SIGKDD explorations newsletter*, 11(1):10–18.

- Hruschka, E. R., Campello, R. J. G. B., Freitas, A., and Carvalho, A. C. P. L. F. (2009). A survey of evolutionary algorithms for clustering. *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews*, 39(2):133–155.
- Kasai, S., Li, F., Shiraishi, J., Li, Q., and Doi, K. (2006). Computerized detection of vertebral compression fractures on lateral chest radiographs: Preliminary results with a tool for early detection of osteoporosis. *Medical Physics*, 33(12):4664–4674.
- Lama, R. S. D. (2018). Uso de redes neurais convolucionais para classificação de fraturas vertebrais por compressão. Undergraduate Thesis, University of São Paulo, Ribeirão Preto, Brazil.
- MacQueen, J. (1967). Some methods for classification and analysis of multivariate observations. In *Proceedings of the 5th Berkeley Symposium on Mathematical Statistics and Probability*, pages 281–297.
- Raineri, L. T. (2018). Seleção de atributos baseada em algoritmos genéticos para o problema de predição de fraturas vertebrais por compressão. Undergraduate Thesis, University of São Paulo, Ribeirão Preto, Brazil.
- Ribeiro, E., Nogueira-Barbosa, M. H., Rangayyan, R., and Azevedo-Marques, P. M. d. (2012). Detection of vertebral compression fractures in lateral lumbar x-ray images. In *XXIII Congresso Brasileiro em Engenharia Biomédica (CBEB)*, pages 1–4.
- Shen, D., Wu, G., and Suk, H.-I. (2017). Deep learning in medical image analysis. *Annual review of biomedical engineering*, 19:221–248.
- Tehranzadeh, J. and Tao, C. (2004). Advances in MR Imaging of Vertebral Collapse. *Seminars in Ultrasound, CT and MRI*, 25(6 SPEC. ISS.):440–460.
- Tinós, R., Zhao, L., Chicano, F., and Whitley, D. (2018). Nk hybrid genetic algorithm for clustering. *IEEE Transactions on Evolutionary Computation*, 22(5):748–761.
- Van Griethuysen, J. J., Fedorov, A., Parmar, C., Hosny, A., Aucoin, N., Narayan, V., Beets-Tan, R. G., Fillion-Robin, J.-C., Pieper, S., and Aerts, H. J. (2017). Computational radiomics system to decode the radiographic phenotype. *Cancer Research*, 77(21):e104–e107.
- Vasconcelos, J. T. S. (2004). *Anatomia Aplicada e Biomecânica da Coluna Vertebral*, volume 2 of *serie*. Etcetera, São Paulo, 2 edition.
- Xu, R., W. D. (2005). Survey of clustering algorithms. *IEEE Transactions on Neural Networks*.
- Xue, B., Zhang, M., Browne, W. N., and Yao, X. (2015). A survey on evolutionary computation approaches to feature selection. *IEEE Transactions on Evolutionary Computation*, 20(4):606–626.