An Approach to Classify Chronic Kidney Diseases using Scintigraphy Images

Elene F. Ohata*‡, Francisco Hércules dos S. Silva*, Shara Shami A. Alves*‡,
Suane Pires P. da Silva*‡, Jefferson S. Almeida* and Pedro Pedrosa R. Filho*‡
*Laboratório de Processamento de Imagens, Sinais e Computação Aplicada (LAPISCO), Instituto Federal do Ceará, Fortaleza, Ceará, Brazil.
Email: {eleneohata, herculessilva, suanepires, jeffersonsilva, shara.shami}@lapisco.ifce.edu.br and pedrosarf@ifce.edu.br
‡Programa de Pós-Graduação em Engenharia de Teleinformática (PPGETI), Universidade Federal do Ceará, Fortaleza, Ceará, Brazil.

***Abstract*—Chronic kidney diseases cause over a million deaths worldwide every year. One of the techniques used to diagnose the diseases is renal scintigraphy. However, the way that is processed can vary depending on hospitals and doctors, compromising the reproducibility of the method. In this context, we propose an approach to process the exam using computer vision and machine learning to classify the stage of chronic kidney disease. An analysis of different features extraction methods, such as Gray-Level Co-occurrence Matrix, Structural Co-occurrence Matrix, Local Binary Patters (LBP), Hu's Moments and Zernike's Moments in combination with machine learning methods, such as Bayes, Multi-layer Perceptron, k-Nearest Neighbors, Random Forest and Support Vector Machines (SVM), was performed. The best result was obtained by combining LBP feature extractor with SVM classifier. This combination achieved accuracy of 92.00% and F1-score of 91.00%, indicating that the proposed method is adequate to classify chronic kidney disease in two stages, being a high risk of developing end-stage renal failure and other outcomes, and otherwise.**

I. Introduction

In 2015, Global Disease Burden Study found that kidney disease is among the top 15 causes of death, comprising more than 1 million deaths worldwide [1]. Chronic Kidney Disease (CKD), also called chronic kidney failure, concerns the loss of the kidney function. In the United States, it is estimated that approximately 30 million adults have CKD [2]. There are 5 stages of CKD and it is measured by observing the Glomerular Filtration Rate (GFR) or kidney damage indicator, or both, during at least three months [3]. GFR number indicates how much of the kidney function a patient has. Lower GFR values means that the kidney disease is getting worse. Table I shows how GFR values relate to CKD stages.

TABLE I: CKD stages based in GRF values.

GFR estimated	CKD stage	% Kidney Function
$>$ **90 ml/min**	1	
61 - 90 ml/min	2	
31 - 60 ml/min	3	
16 - 30 ml/min	4	
$\leq$ **15 ml/min**	5	

Hypertension, metabolic syndrome, and diabetes are the primary diseases that can increase the risk of an adult to develop CKD [4]. Furthermore, a person may have aggravations due to the CKD, such as cancer, bone disease, anemia and increased risk of cardiovascular disease [5]. The early detection of CKD might prevent the aforementioned CKD aggravations as well as avoid further health problems by means of medications and a healthier lifestyle [2].

Regarding the estimation of GFR followed by the detection of CKD, there are two main methods for this purpose: (i) the measurement of creatine elimination which takes blood and urine exams over a 24 hour period to compare the creatine levels in both exams, and (ii) the dynamic renal scintigraphy [6]. One advantage of the scintigraphy is that it can show the corporal activity in regions of interest [7].

Gates' method is the standard method applied to estimate the GFR in scintigraphy exams [8]. First of all, we have the dynamic renal scintigraphy images, then the kidney and the background radiation regions are manually marked. After that, the renal activity is measured using the marked regions by counting the intensity of the pixels in each one of the images for each kidney from the original scintigraphy images. Finally, GFR is estimated using the Gates' algorithm which is calculated employed the value of the intensity of the pixels [9]. Those pixels intensity counts generate a curve of renal activity versus time, which is called renogram.

The process of manually marking the regions of interest (ROI) is time consuming, tedious and depends on the specialist skill. Consequently, the results are dependent on the person processing the exam, which can compromise the reproducibility of the method [10]. Thus, methods have been proposed to support the definition of ROIs. Nonetheless, those techniques still involve the process of counting the intensity of the pixels to estimate the GFR, added with the segmentation error. In [11], the authors applied adaptive edge-based active contour techniques to automatically segment the kidney in scintigraphy

exams. However, the technique requires that the operator provides a starting point for the segmentation algorithm. In [12], another semi-automatic method for segmentation of renal scintigraphy images using gradient map, adaptive threshold and region growing was proposed. Nevertheless, it is necessary an input provided by the operator in order to correctly segment the kidneys. In [10], their automatic segmentation approach could not segment a patient kidney with poor renal function.

In this paper, we propose a new approach for diagnosis support of CKD classes from scintigraphy images using computer vision, with a focus on the severity of the CKD. Our approach considers two classes, being: (i) higher risk of developing kidney failure, and (ii) lower risk. To the best of our knowledge, this is the first attempt to aid in CKD diagnosis with neither need for segmentation methods nor input of operators.

To evaluate the proposed method, an analysis was made of several combinations of features extractor and classifiers. The feature extractors considered were Gray-Level Co-occurrence Matrix (GLCM), Local Binary Patterns (LBP), Hu's Moments (HM), Structural Co-occurrence Matrix (SCM) and Zernike's Moments. The classifiers considered were Bayes, Multi-layer Perceptron (MLP), k-Nearest Neighbors (kNN), Random Forest (RF) and Support Vector Machines (SVM).

This paper is organized as follows: Section II presents our methodology in details. Section III shows the experimental results and discussion, followed by the conclusion and future works in Section IV.

II. Methodology

The proposed approach flowchart is depicted in Figure 1 and it consists of 5 stages explained in details in the next sections. In summary, in the first stage, we have the exam images of the patients as input. Then, those images go through the pre-processing stage resulting in a reduced set of 15 images for each exam. After that, the attributes vectors are generated during the feature extraction stage. Finally, those vectors are classified in Class 0 or 1 by machine learning methods.

This study aims at aiding in the diagnosis of the CKD stage through the classification of scintigraphy images of a patients' exam to determine the need for specialized monitoring.

A. Acquiring images

The images used in this study are from the Database of Dynamic Renal Scintigraphy [13]. The training dataset is from the *drsprg* dataset, which contains scintigraphy images of 107 patients' exams previously classified in one of the five CKD stages, depending on GFR estimation from collected blood samples.

In *drsprg* dataset, each exam took 30 minutes and the stored exam consists of 180 images of size 128×128 pixels that were captured every 10 seconds over the exam procedure. Both anterior and posterior projections were captured as well. We noticed that 5 patient's exams do not have all the 180 images, to wit, patients numbered 25 (CKD 1), 55 (CKD 2), 65 (CKD 2), 81 (CKD 3) and, 106 (CKD 5); those ones were removed from the experiment in order to keep it standardized.

A patient classified in CKD 3 or above has higher risks of mortality [14] or other outcomes, *i.e.* bone disease, and the patients from this group have more sensibility to side effects from drugs [15]. Therefore, we have considered working with two classes. The Class 0 comprehends the CKD stages 1 and 2, and the Class 1 the remaining CKD stages: 3, 4 and 5. Thus, Class 1 indicates whether the patient needs specialized monitoring due to further disease progressions, and Class 0, otherwise. Hence, we have 44 samples from Class 0, and 58 from Class 1.

B. Pre-processing

All 180 images of each patient's exam have been blurred through the application of a mean filter. After that, the set of 180 images was divided into 15 subsets, then, each subset represents a 2-minute frame taken from the exam, and contains 12 images each. Finally, an average image of each one of those 15 subsets was generated. The average image is the sum of a subset of 12 images, and divided by the total of images (12). The 2-minutes period that was chosen is a common practice in the literature, one can see details in [8], [12].

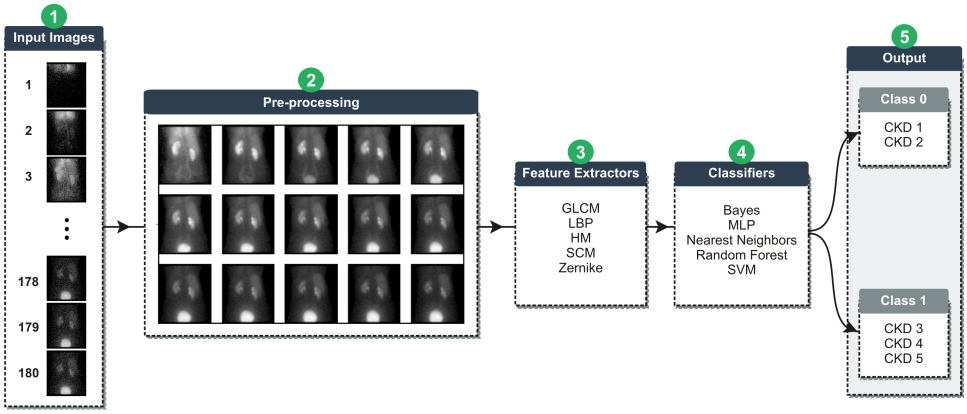


Fig. 1: Methodology of the proposed approach for Chronic Kidney Disease classification.

C. Feature extraction

The goal in feature extraction techniques is obtaining attributes and characteristics from an input image. Those methods work on different perspectives of the image primarily on moments and texture. For this work, we applied the feature extractors: GLCM [16], LBP [17], HM [18], Zernike's Moments [19] and SCM [20], which return 9, 108, 7, 6 and 8 attributes, respectively.

For each feature extraction method, each patient resulted in 15 vectors of features. Those vectors were concatenated, so that a patient is represented by one vector. For example, SCM extracts 8 features per image so the patient's representing vector comprehends 120 features.

The proposed approach uses all the mentioned feature extractors to obtain different significant characteristics that may facilitate the classification among CKDs without the need for the segmentation of scintigraphic exams.

D. Classification

We adopted different classifiers in order to find the best configuration for supporting the diagnosis of CKD. Then, we chose classifiers that are from different types and self-configurable. From the concept of probability and statistics, we use the Bayes classifier [21]. MLP [22] is an architecture of artificial neural network. KNN [23] is based on instances, while RF [24] is based on the Decision Tree Method. The SVM classifier [25] is based on the Statistical Learning Theory.

In regards to the classifier's parameters, they were selected through the random search algorithm on hyperparameters configurations for the classifiers aiming to obtain the best performances. For SVM, the kernels linear, polynominal and radial basis function (RBF) were considered. After the hyperparameters optimization, 10-fold cross-validation was done. The number $k = 10$ was chosen so that there were more samples, bringing the model closer to reality. All the experiments were performed on the training dataset.

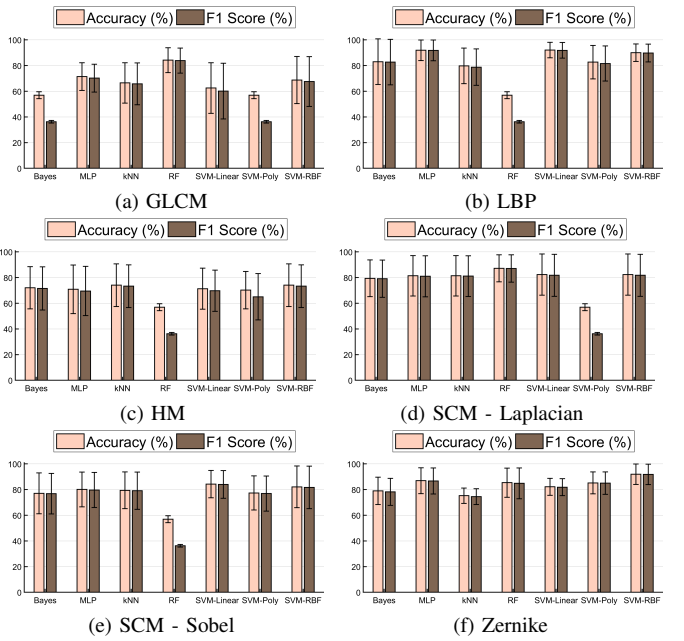
E. Evaluation Metrics

In order to evaluate the classifiers, four evaluation metrics were used: Accuracy (Acc), Precision (Pre), F1-Score (F1) and Recall (Rec), which are obtained by means of the confusion matrix, where True Positives (TP) represents the number of times that the classifier predicts a patient that has a lower risk of mortality, *i.e.* Class 0. False Negatives (FN) informs the total of misclassified patients as high CKD stage, while False Positives (FP) indicates how many patients in Class 1 were classified as Class 0. Finally, True Negatives (TN) presents the number of patients of Class 1 that are predicted correctly.

III. Results

In this session, we compare the achieved results by the combination of different feature extractors and classifiers considered in our approach. The tests were performed on an iMac with Intel Core i5 processor, running at 3.2 GHz, and 16GB RAM.

Figure 2 graphically presents the classification results, consisting of mean value and standard deviation, for all combinations of classifiers and extractors. Accuracy and F1-Score metrics express the database's similarity to the number of false-negative and false-positive results. Looking at Figure 2, we conclude that both metrics presented similar values in most combinations plotted. MLP and SVM classifiers obtained the best performances when combined with LBP, as seen in Figure 2b, and with Zernike, as shown in Figure 2f.



(a) GLCM (b) LBP (c) HM (d) SCM - Laplacian (e) SCM - Sobel (f) Zernike

Fig. 2: Accuracy and F1-Score for each combination of feature extractor with classifier.

Table II shows average value and standard deviation of Accuracy, Precision, F1-Score, Recall, training and testing times from the best combinations of feature extractors with classifiers. The greatest values of Accuracy and F1-Score were reached by the combination of LBP with SVM-Linear, obtaining 92.05% and 91.89%, respectively. While the combination of Zernike with SVM-RBF got the best values of Precision and Recall, achieving 93.40% and 92.42%, respectively.

According to Table II, the combination of Zernike's method with SVM-RBF achieved the shortest training time, which was an average of 0.005s. MLP classifier obtained the highest training time, reaching 12.356s, when combined with LBP.

In addition, regarding the same data, SVM-RBF in combination with Zernike was the fastest, performing this step in 0.003s. MLP classifier was the slowest classifier in the classification step when combined with LBP, returning a testing time of 0.011s.

IV. Conclusion and Future Work

In this article, a new approach to aid in the diagnostic of chronic kidney disease using scintigraphy images is proposed. In addition, an analysis was performed combining several types of classifiers and feature extractors.

TABLE II: Acc, Pre, F1, Rec, training and testing times obtained by the best combinations of features extractors and classifiers.

Feature extractor	Classifier	Setup	Acc (%)	Pre (%)	F1 (%)	Rec (%)	Training Time (s)	Testing Time (s)
LBP	**SVM**	**Linear**	**92.05 ±6.04**	**92.74 ±5.88**	**91.89 ±6.13**	**92.25 ±6.06**	**0.031 ±0.001**	**0.005 ±0.000**
Zernike	SVM	RBF	91.96 ±7.93	93.40 ±6.60	91.84 ±7.96	92.42 ±6.67	0.005 ±0.001	0.003 ±0.000
LBP	MLP	-	91.94 ±8.03	92.20 ±8.06	91.85 ±8.12	92.25 ±8.03	12.356 ±0.544	0.011 ±0.003
LBP	SVM	RBF	90.03 ±6.82	90.77 ±7.03	89.81 ±6.90	90.00 ±6.97	0.031 ±0.001	0.006 ±0.000

According to the results shown, we can conclude that the combination of LBP with SVM-Linear is suitable to classify the CKD in two classes. This combination reaches 92.00% in Accuracy, 92.00% in Precision, 91.00% in F1-score and 92.00% in Recall, showing that is possible to classify CKD stage using only images, without the necessity of manually marking the kidneys region. Then, the patients can be classified in higher risk of mortality and kidney failure, or lower risk of aggravation of the kidney function.

For future improvements, first, we intend to apply statistical tests. Secondly, we aim to expand the study to the original 5 CKD stages. Besides, since Zernike and LBP brought the best results, a combination of both might be worthwhile to enhance the results, making use of Zernike's advantage to discriminate shapes, and LBP's capability to process textures. Furthermore, Convolutional Neural Networks (CNN) can be applied as feature extractors with several architectures, such as VGG16 [26] and MobileNet [27], and other machines learning methods can be used, such as Optimum-Path Forest.

ACKNOWLEDGMENT

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001". Also Pedro Pedrosa Rebouças Filho acknowledges the sponsorship from the Brazilian National Council for Research and Development (CNPq) via Grants Nos. 431709/2018-1 and 311973/2018-3.

REFERENCES

[1] T. Vos, C. Allen, M. Arora, R. M. Barber, Z. A. Bhutta, A. Brown, A. Carter, D. C. Casey, F. J. Charlson, A. Z. Chen *et al.*, "Global, regional, and national incidence, prevalence, and years lived with disability for 310 diseases and injuries, 1990–2015: a systematic analysis for the global burden of disease study 2015," *The Lancet*, vol. 388, no. 10053, pp. 1545–1602, 2016.

[2] C. for Disease Control and Prevention, "National chronic kidney disease fact sheet, 2017," *Atlanta, GA: US Department of Health and Human Services, Centers for Disease Control and Prevention*, 2017.

[3] K. Larmour, A. Maxwell, and A. Courtney, "Improving early detection of chronic kidney disease." *The Practitioner*, vol. 259, no. 1779, pp. 19–23, 2015.

[4] K.-U. Eckardt, J. Coresh, O. Devuyst, R. J. Johnson, A. Köttgen, A. S. Levey, and A. Levin, "Evolving importance of kidney disease: from subspecialty to global health burden," *The Lancet*, vol. 382, no. 9887, pp. 158–169, 2013.

[5] A. C. Webster, E. V. Nagler, R. L. Morton, and P. Masson, "Chronic kidney disease," *The Lancet*, vol. 389, no. 10075, pp. 1238–1252, 2017.

[6] A. Taylor Jr and J. Nally, "Clinical applications of renal scintigraphy." *AJR. American journal of roentgenology*, vol. 164, no. 1, pp. 31–41, 1995.

[7] M. E. Kerl and C. R. Cook, "Glomerular filtration rate and renal scintigraphy," *Clinical techniques in small animal practice*, vol. 20, no. 1, pp. 31–38, 2005.

[8] W. H. S. dos Santos, S. T. M. Ataky, A. C. Silva, A. C. de Paiva, and M. Gattass, "Automatic method for quantitative automatic evaluation in dynamic renal scintilography images," *Multimedia Tools and Applications*, vol. 76, no. 18, pp. 19 291–19 315, 2017.

[9] Y. Qi, P. Hu, Y. Xie, K. Wei, M. Jin, G. Ma, Q. Li, B. Xu, and X. Chen, "Glomerular filtration rate measured by 99mtc-dtpa renal dynamic imaging is significantly lower than that estimated by the ckd-epi equation in horseshoe kidney patients," *Nephrology*, vol. 21, no. 6, pp. 499–505, 2016.

[10] K.-J. Lin, J.-Y. Huang, and Y.-S. Chen, "Fully automatic region of interest selection in glomerular filtration rate estimation from 99mtc-dtpa renogram," *Journal of digital imaging*, vol. 24, no. 6, pp. 1010–1023, 2011.

[11] P. Suapang, K. Dejhan, and S. Yimman, "The estimation of gfr and erpf using adaptive edge-based active contour for the segmentation of structures in dynamic renal scintigraphy," *International Journal of Innovative Computing, Information and Control*, vol. 11, no. 1, pp. 87–103, 2015.

[12] Y. Aribi, F. Hamza, W. Ali, A. M. Alimi, and F. Guermazi, "An automated system for the segmentation of dynamic scintigraphic images," *Applied Medical Informatics*, vol. 34, no. 2, pp. 1–12, 2014.

[13] "Database of dynamic renal scintigraphy," [Online; accessed: 12-march-2019]. [Online]. Available: http://www.dynamicrenalstudy.org/

[14] A. S. Levey, J. Coresh, E. Balk, A. T. Kausz, A. Levin, M. W. Steffes, R. J. Hogg, R. D. Perrone, J. Lau, and G. Eknoyan, "National kidney foundation practice guidelines for chronic kidney disease: evaluation, classification, and stratification," *Annals of internal medicine*, vol. 139, no. 2, pp. 137–147, 2003.

[15] K. Harris and B. Stribling, "Automated estimated gfr reporting: A new tool to promote safer prescribing in patients with chronic kidney disease?" *Therapeutics and clinical risk management*, vol. 3, no. 5, p. 969, 2007.

[16] R. M. Haralick, K. S. Shanmugam, and I. Dinstein, "Textural features for image classification," vol. 3, no. 6, 1973, pp. 610–621.

[17] T. Ojala, M. Pietikäinen, and T. Mäenpää, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 7, pp. 971–987, 2002.

[18] M.-K. Hu, "Visual pattern recognition by moment invariants," *IRE transactions on information theory*, vol. 8, no. 2, pp. 179–187, 1962.

[19] A. Khotanzad and Y. H. Hong, "Invariant image recognition by zernike moments," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 12, no. 5, pp. 489–497, 1990.

[20] G. L. B. Ramalho *et al.*, "Rotation-invariant feature extraction using a structural co-occurrence matrix," *Measurement*, vol. 94, no. 2, pp. 406–415, 2016.

[21] S. Theodoridis and K. Koutroumbas, *Pattern Recognition (Fourth Edition)*. USA: Academic Press, 2008.

[22] S. Haykin, *Neural Networks and Learning Machines*. McMaster University, Canada: Prentice Hall, 2008.

[23] K. Fukunaga and P. M. Narendra, "A branch and bound algorithm for computing k-nearest neighbors," *IEEE Transactions on Computers*, vol. C-24, no. 7, pp. 750–753, July 1975.

[24] L. Breiman, "Random forests," *Machine learning*, vol. 45, no. 1, pp. 5–32, 2001.

[25] V. N. Vapnik, *Statistical Learning Theory*. Nova Jersey, EUA: John Wiley & Sons, 1998.

[26] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.

[27] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam, "Mobilenets: Efficient convolutional neural networks for mobile vision applications," *arXiv preprint arXiv:1704.04861*, 2017.