Autoregulatory Efficiency Assessment in Kidneys Using Deep Learning

Sebastian Alphonse
Dept. of Elec. and Comp. Engr.
Illinois Institute of Technology
Chicago, IL, U.S.A.
salphons@hawk.iit.edu

Aaron J. Polichnowski
Department of Biomedical Sciences
East Tennessee State University
Johnson City, TN, U.S.A.
polichnowski@mail.etsu.edu

Karen A. Griffin, Anil K. Bidani
Departments of Medicine
Loyola Univ. Med. Ctr. and *Edward Hines, Jr. VA Hosp.*
Maywood, IL, U.S.A.
{karen.griffin2,anil.bidani}@va.gov

Geoffrey A. Williamson
Dept. of Elec. and Comp. Engr.
Illinois Institute of Technology
Chicago, IL, U.S.A.
williamson@iit.edu

***Abstract*—A convolutional deep neural network is employed to assess renal autoregulation using time series of arterial blood pressure and blood flow rate measurements in conscious rats. The network is trained using representative data samples from rats with intact autoregulation and rats whose autoregulation is impaired by the calcium channel blocker amlodipine. Network performance is evaluated using test data of the types used for training, but also with data from other models for autoregulatory impairment, including different calcium channel blockers and also renal mass reduction. The network is shown to provide effective classification for impairments from calcium channel blockers. However, the assessment of autoregulation when impaired by renal mass reduction was not as clear, evidencing a different signature in the hemodynamic data for that impairment model. When calcium channel blockers were given to those animals, however, the classification again was effective.**

***Index Terms*—machine learning, neural networks, biomedical signal processing, physiology, nephrology**

I. Introduction

Chronic kidney disease (CKD) is a growing health concern, affecting approximately 13.4% of the world population [1]. Hypertension plays a major role in the progression of CKD to End Stage Renal Disease (ESRD) [2]–[4], and renal blood flow (RBF) autoregulation (AR) efficiency appears to be a major determinant of the susceptibility to hypertensive renal injury [3], [5]. Assessment of AR efficiency is therefore a useful diagnostic tool.

For assessing AR efficiency in animal models of CKD, the most widely accepted methodology is the measurement of steady-state responses in RBF to acute, large changes in arterial blood pressure (BP) mediated by aortic clamps with the animal under anesthesia [6]–[9]. The average BP and RBF values before and after the acute BP changes are used to calculate an Autoregulatory Index (ARI). This approach suffers not only from the impact of anesthesia on physiological quantities associated with AR [10], [11] but also from its neglect of the kinetics of the AR response, which is an important factor affecting the impact of AR impairment on CKD progression [6], [12], [13]. Even when kinetics are characterized via these "step" AR studies [9], the impact of anesthesia remains.

Alternatively, one may assess AR efficiency using BP and RBF data collected in the conscious state where spontaneous fluctuations in BP initiate AR responses manifested in the RBF. A common approach to AR assessment in the conscious state involves empirical transfer function estimation representing the dynamics between variational BP and variational RBF [6], [8], [9]. However, estimates of AR efficiency using these approaches neither yield consistent estimates nor correlate well with the susceptibility to renal hypertensive injury [6]. More promising is a methodology that uses the same kind of ARI calculation employed in the steady-state "step" studies but applies it to pressure and flow values for adjacent short segments within recordings between which the average pressure undergoes a sizable change [13]. These Short Segment ARI (SSARI) values provide improved assessment of AR efficiency when a sufficient number of qualifying pressure change events enable averaging to smooth the inherent variability in the SSARI.

Our work here considers a deep learning approach to evaluating AR efficiency. We develop a Deep Neural Network (DNN) to classify or score renal hemodynamic pressure and flow signals on a scale from "intact AR" to "impaired AR." The DNN in such an approach selects on its own the relevant features to extract from the data in order to perform the

This work was supported by National Institutes of Diabetes and Digestive and Kidney Diseases grants DK-40426 (to Dr. Bidani) and DK-61653 (to Dr. Griffin), by a Merit Review Award (to Dr. Griffin) and Career Development Award (1IK2BX001285 to Dr. Polichnowski) from the Office of Research and Development of the Department of Veterans Affairs, and a Career Development Award from the National Kidney Foundation of Illinois (to Dr. Polichnowski). Dr. Polichnowski is currently supported by a Carl Gottschalk Research Scholar Grant from the American Society of Nephrology Foundation for Kidney Research and the American Heart Association (17AIREA33660433).

classification or scoring, rather than being forced to use preselected features. Having said this, we note that we mask the average BP and RBF values, normalizing each to unit value, so that they may not be used to classify or score the data. This is important in part because animal models for impaired AR may involve coincidental changes in RBF. For instance, this is the case with reduction of the renal mass: removal of one kidney and surgical excision of two-thirds of the other. In an animal with reduced renal mass, the RBF level in the remaining (partial) kidney is approximately double the value that would be obtained without renal mass reduction. We also normalize the signal variance to unit value.

To examine the viability of DNNs for assessing AR efficiency, we train a DNN to differentiate between BP and RBF data from rats with intact AR and from rats with a particular type of AR impairment. We then present these DNNs with BP and RBF data from rats with different AR impairment. The DNN takes as input one minute long "snippets" of both BP and RBF data and scores that snippet on a scale from zero (intact) to one (impaired). An average score across all snippets for a single rat comprises the assessment for that animal. No single quantitative evaluation of the achieved results is by itself fully informative, but qualitatively we are able to see the potential of deep learning to assess AR efficiency using a particular quantitative metric separating intact from impaired AR.

II. Data Collection

Our development of DNNs to assess renal AR efficiency has been performed to date with respect to two different rodent models for impaired AR as well as with respect to normal, intact AR. In each case, data were collected in Sprague-Dawley rats instrumented to allow chronic monitoring of BP and RBF in a conscious, unrestrained state. Rats had a BP sensor inserted into the aorta below the renal artery and a radiotransmitter fixed to the peritoneum as previously described [6], [14], [18], [19]. Rats were also equipped with an ultrasonic transit time flow probe around the renal artery, with the probe cable affixed to the back muscles and exiting the animal at the neck as previously described [6], [14], [20], [21]. Following recovery from instrumentation, simultaneous recordings of BP and RBF were taken using a sampling rate of 200 Hz for periods of one to two hours.

Data collected from rats as described above but otherwise with no additional interventions are associated with normal, intact AR. The two models used to represent impaired AR were that stemming from renal mass reduction (RMR) and that stemming from administration of a calcium channel blocker (CCB). In the case of RMR, the rat underwent removal of the right kidney (right uninephrectomy) and surgical excision of both poles of the left kidney, referred to as RK-NX [22], [23]. Recordings of BP and RBF were made about three or four weeks after RMR surgery. In the case of CCBs, some rats received the CCB amlodipine in their drinking water at concentrations of either 100 mg/l or 200 mg/l [14]. After 48 to 72 hours of receipt of the drug, recordings of BP and RBF were taken. Other rats received the CCB mibefradil via standard rodent chow formulated to contain mibefradil (either 0.065% or 0.1%) [14]. BP and RBF recordings were taken subsequent to 48 to 78 hours from the initiation of the mibefradil diet.

In all situations, from one to four recordings of BP and RBF, each of length one to two hours, were taken for each rat. The number of rats in each category is shown in Table I. Within each of these recordings, a 30 minute segment was selected according to criteria for ensuring good quality data. First, consideration was restricted to sections of the recording without strong noise or other artifacts. Second, cross spectral analysis was performed on each available artifact-free 30 min. subsegment. The average coherence in the frequency band from 0.3 Hz to 0.7 Hz was calculated for each such subsegment, and the 30 min. of data chosen for analysis was that which maximized this average coherance. The 0.3 Hz to 0.7 Hz frequency range lies just above the area where the characteristic frequency associated with the myogenic component of the AR response is expected. Higher coherence values in this range indicate data for which a linear analysis better represents the dynamic AR response. The spectral analysis is performed using Welch's averaged periodogram method as described in [14] using 50% overlapping sections of 81920 samples (approximately 7 min.) to which Hanning windows were applied.

TABLE I
Number of Rats

Recording type	Number
Intact	230
Amlodipine 100 mg/l	27
Amlodipine 200 mg/l	5
Mibefradil 0.065% or 0.1%	8
RK-NK	76
RK-NX + Amlodipine	21

III. Preparation of the Data

We separate the selected 30 minute segments of data for each rat into one minute long snippets, resulting in 30 to 120 snippets for each rat, depending on the number of recordings. For training data, additional diversity is achieved by randomly selecting the starting point of the first snippet within the first minute of a given 30 minute segment, making 29 snippets available for use from that segment. Periodically during training, this starting point is randomized again. For test data, we simply use the 30 one minute snippets that occur end to end within the full 30 minute segment.

Prior to input to the DNN, each one minute snippet is normalized to have zero mean and unit variance. We perform the normalization of the mean value in order to prevent the network from using mean BP or RBF in its assessment. Certain types of animal models of impaired AR, such as RMR, can impact baseline RBF means, while AR action is reflected directly in relative values of RBF following changes in BP.

Removing means will not affect the network's capacity to observe relative changes in value or dynamics of values, but does prevent the secondary characteristic of baseline BP or RBF from influencing assessment. Similarly, normalizing the variance inhibits the network from utilizing basic variability of data in its AR assessment.

IV. Network Architecture

The DNN that we designed for classifying the intact and impaired AR is diagrammed in Fig. 1. We use a standard structure consisting of two convolutional layers followed by two fully connected layers. This basic architecture has been useful in applications similar to ours [15]–[17]. The DNN takes as input one minute long snippets of BP and RBF data. The first convolutional layer has 16 filters each with length of 2 seconds (400 samples). These filters operate on both the BP and RBF data and learn the temporal relationship between these two signals, which is important for learning the AR characteristics. These filters move with a stride of one sample each yielding a output vector of size 11601×1. This is followed by the max pooling layer that operates on 10×1 vectors and moves with a stride of 2, producing a 5799×1 output vector for each filter. The second convolutional layer has 16 filters, each operating on the first max pooling layer output and moving with a stride of one sample to yield the $5700 \times 1 \times 16$ output. The second convolutional layer is followed by a max pooling layer of shape 8×1 and moves with a stride of 2 yielding the output shape of $2848 \times 1 \times 16$. The output of the second max pooling layer is flattened and fed as input to the fully connected layers. The first fully connected layer has 256 units and the second fully connected layer has 128 units. In all these hidden layers, a Rectified Linear Unit (Relu) is used as the activation function. The output of the second fully connected layer is fed to the output layer, which uses a sigmoid activation function. The output of the sigmoid activation function is between 0 (corresponding to intact AR) and 1 (corresponding to impaired AR).

This network architecture was arrived at after much experimentation. In addition to trying just convolutional layers and just fully connected layers, we varied the number and type of layers, size of input data, number and length of filters in the convolutional layer, size of the fully-connected layers, and other network parameters. The architecture presented here is the best performing of all these variations.

V. Training and Testing

We train the network to distinguish between BP/RBF data from the intact AR group (230 rats) and that receiving the CCB amlodipine at the 100 mg/l level (27 rats). We use 60% of the rats for training, 6% for validation, and 34% for testing. This results in 138 intact and 16 impaired rats available for training. To balance the amount of data during training, we initially select 16 intact rats at random as the source of training data, and after every five epochs of training we reselect 16 intact rats, again at random, to provide data for the next five epochs.

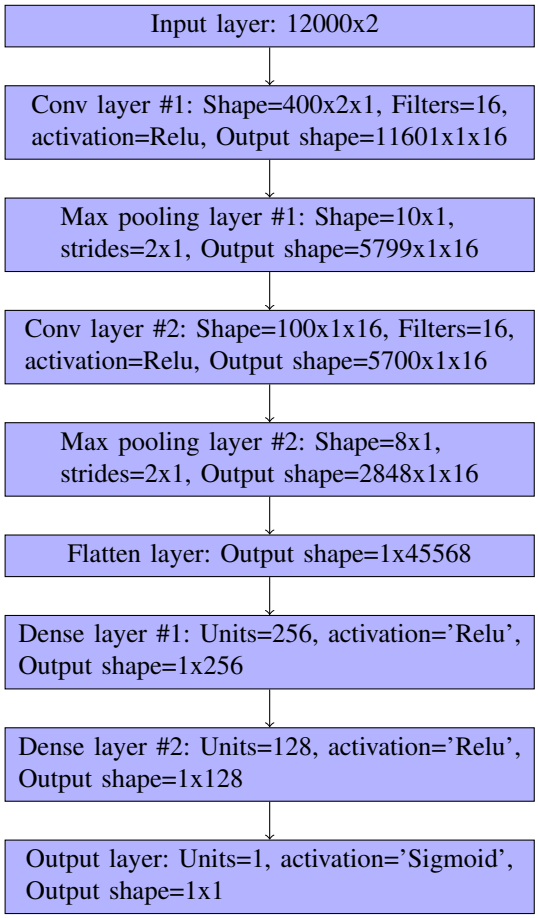

Fig. 1. *Deep Neural Network architecture used for classifying intact vs impaired BP/RBF data snippets. The input (12000×2) data comes from one minute of BP and RBF data sampled at 200 Hz. The output of the network is from sigmoid activation units taking values between zero and one. Score zero indicates intact AR, and score one indicates AR impairment.*

Recall from Section III that each rat used for training data has one or more 30 minute recordings of BP and RBF data, with each recording's data split into successive one minute snippets starting from a randomly chosen and periodically reset sample within the first minute of data. This randomization is a kind of data augmentation technique frequently used with a relatively small set of training examples. We trained the networks through multiple epochs to classify the data on an AR scale from intact (score 0) to impaired (score 1). We reduce the learning rate by 2 as the loss plateaus. We used an Adam optimizer [24] to reduce the binary cross entropy loss between the predicted output class and the true class.

Following training, the test data was used in evaluation of network performance. All available recordings for each rat in the test group are used to generate the maximum number of one minute long input snippets. Each of these is presented to the network, resulting in an output score in the range from 0 to 1. Scores for all snippets are averaged to produce the score for that rat. With this type of training scheme, we trained five models with different sets of training and test data partitions. The results obtained from these models were

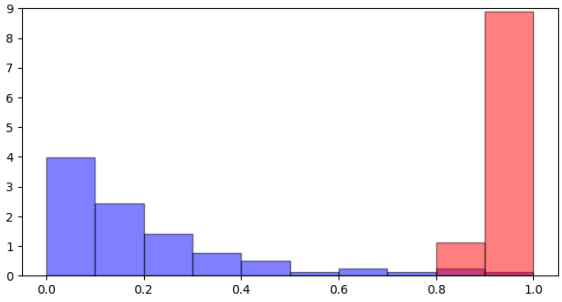

Fig. 2. *Score pdf for test data from rats with intact AR and rats receiving amlodipine, concentration 100 mg/l. The network was trained to distinguish betweeen training data from these two groups. Intact rats (blue) number 78, with 71 scoring below 0.5. Amlodipine 100 rats (red) number nine, with all nine scoring above 0.5.*

quantitatively similar. This similarity suggests that the DNN is not overfitting, indicating that our augmented dataset is sufficient for training the models.

We also apply as input to the network the BP/RBF data from a selection of other populations of data sets associated with impaired AR. There are four such groups: CCB amlodipine at a higher concentration of 200 mg/l, CCB mibefradil (combining two fraction levels of 0.065% and 0.1%), the RMR model RK-NX, and RK-NX rats also receiving the CCB amlodipine. All data from these groups are presented to the network to test its performance in assessing AR efficiency.

VI. Results and Discussion

The results from the various experiments are presented as empirical probability density functions (pdfs) or histograms. We quantize to score intervals of width 0.1, so that we have 10 bins in the range $[0, 1]$ into which all average scores may fall. The pdfs are shown in Figs. 2 through 6. We emphasize score pdfs over confusion matrices or receiver operating characteristics (ROCs) because we are interested in the DNN output score as an assessment measure for autoregulatory efficiency rather than only to make a classification decision. Nonetheless, to convey the classification potential, we also present in Fig. 7 the ROC for the intact AR and amlodipine test data. In discussing the score pdfs, we use as an aggregated indication of overall general performance the fraction of test rats scoring below and above the midpoint value of 0.5, with the range $[0, 0.5]$ representing an intact AR score, and the range $[0.5, 1]$ representing an impaired AR score.

We see in Fig. 2 that the two groups visually separate into the proper halves of the score range, with 87% of the testing data scoring in the appropriate half of the range. From a classification perspective, this 87% performance stemming from an arbitrary threshold choice of 0.5 in score may be interpreted as a good result. Indeed, in a study in which 179 classifiers from 17 distinct classifier families were applied to 121 different data sets, the average best performance on the data sets was 86.9% [25]. That the rats receiving amlodipine have scores concentrated at the high end of the range enables even better classification performance, as shown in the ROC

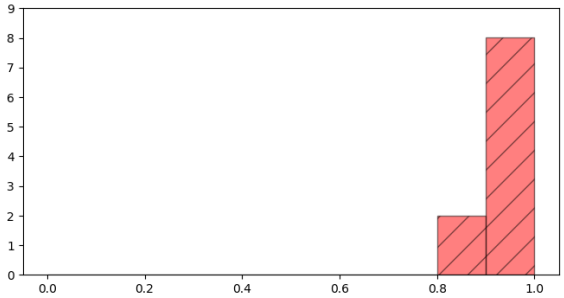

Fig. 3. *Score pdf for test data from rats receiving amlodipine, concentration 200 mg/l. Number of rats is five, with all five scoring above 0.5.*

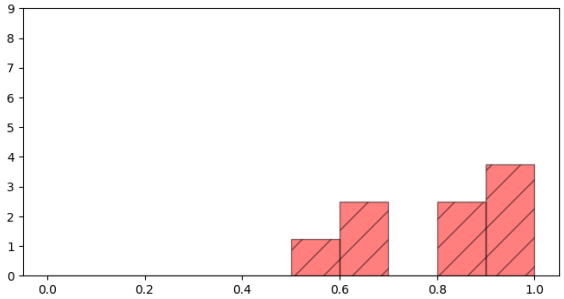

Fig. 4. *Score pdf for test data from rats receiving mibefradil at two fractions (0.065% and 0.1%). Number of rats is eight, with all eight scoring above 0.5.*

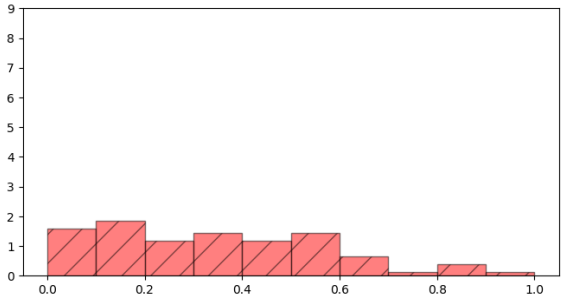

Fig. 5. *Score pdf for test data from the RK-NX group. Number of rats is 76, with 21 scoring above 0.5 and 55 scoring below 0.5.*

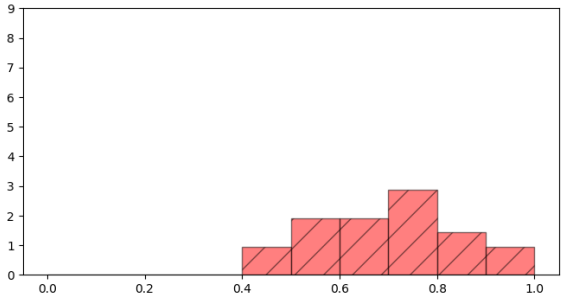

Fig. 6. *Score pdf for test data from RK-NX rats also receiving amlodipine. Number of rats is 21, with 19 scoring above 0.5 and two scoring below 0.5.*

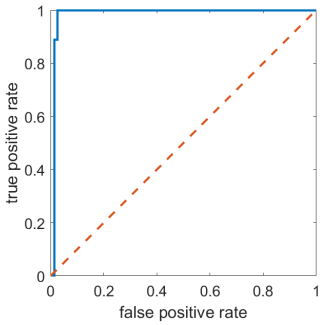

Fig. 7. *ROC for test data from rats with intact AR and rats receiving amlodipine, concentration 100 mg/l. The area under the ROC is 0.9858.*

of Fig. 7, whose area under the curve is 0.9858. These results give confidence that the kind of data with which we are working allows for reasonable discrimination between intact and impaired AR of this type. To the authors' knowledge, this DNN represents the first classification scheme for distinguishing between intact and impaired AR using BP and RBF measurements acquired in the conscious state.

Figs. 3 to 6 show the network's performance when presented with data from models of AR impairment distinct from what was used in training. In Fig. 3, we see higher scores, corresponding to a stronger indication of AR impairment, in response to BP-RBF data from rats given a higher dose of amlodipine. Fig. 4 shows scores for data from rats given mibefradil, a different CCB. In both these cases, 100% of the rats scored in the impaired range (score > 0.5). For a different model of AR impairment, that of RK-NX RMR, the results appear in Fig. 5. Here, scores are spread across the whole range of values, showing that the signature for RK-NX differs from that for CCBs. However, giving RK-NX rats amlodipine restores the capacity of the network to classify correctly the impairment, with Fig. 6 showing a return of scores to the "impaired" half of the range.

VII. Conclusion

These results show the promise of DNNs to classify and assess renal AR efficiency from BP-RBF data. Training based on one CCB model of impairment appears to have capability to assess impairment from other CCBs and also CCBs in the presence of RMR. However, we view a conclusion that the score reflects the extent of impairment to be premature, given the similarity of scores for two dose levels of amlodipine. Furthermore, the CCB-trained DNN does not pick up AR impairment from RMR alone. Future work will seek to improve DNN performance with respect to these and other factors.

References

[1] N.R. Hill, S.T. Fatoba, J.L. Oke, J.A. Hirst, C.A. O'Callaghan, D.S. Lasserson, and F.D. Richard Hobbs, "Global prevalence of chronic kidney disease - a systematic review and meta-analysis," *PLoS One*, vol. 11, no. 7, July 2016.

[2] P. Rossing, E. Hommel, U.M. Smidt, and H.H. Parving, "Impact of arterial blood pressure and albuminuria on the progression of diabetic nephropathy in IDDM patients," *Diabetes*, vol. 42, pp. 715-719, 1993.

[3] A.K. Bidani, A.J. Polichnowski, R. Loutzenhiser, and K.A. Griffin, "Renal microvascular dysfunction, hypertension and CKD progression," *Curr Opin Nephrol Hypertens*, vol. 22, pp. 1-9, 2013.

[4] R. Saran et al., "US renal data system 2018 annual data report: epidemiology of kidney disease in the United States," *Am J Kidney Dis*, vol. 73, pp. A7-A8, 2019.

[5] A.K. Bidani and K.A. Griffin, "Pathophysiology of hypertensive renal damage: implications for therapy," *Hypertension*, vol. 44, pp. 595-601, 2004.

[6] A.K. Bidani, R. Hacioglu, I. Abu-Amarah, G.A. Williamson, R. Loutzenhiser, and K.A. Griffin, , " 'Step' vs. 'dynamic' autoregulation: implications for susceptibility to hypertensive injury," *Am J Physiol Renal Physiol*, vol. 285, pp. F113-120, 2003.

[7] M. Carlstrom, C.S. Wilcox, and W.J. Arendshorst, "Renal autoregulation in health and disease," *Physiol Rev*, vol. 95, pp. 405-511, 2015.

[8] W.A. Cupples and B. Braam, "Assessment of renal autoregulation," *Am J Physiol Renal Physiol*, vol. 292, pp. F1105-1123, 2007.

[9] A. Just, "Mechanisms of renal bllod flow autoregulation: dynamics and contributions," *Am J Physiol Regul Integr Comp Physiol*, vol. 292, pp. R1-17, 2007.

[10] L.G. Navar, E.W. Inscho, S.A. Majid, J.D. Imig, L.M. Harrison-Bernard, and K.D. Mitchell, "Paracrine regulation of the renal microcirculation," *Physiol Rev*, vol. 76, pp. 425-536, 1996.

[11] W.A. Pettinger, "Anesthetics and the renin-angiotensin-aldosterone axis," *Anesthesiology*, vol. 48, pp. 393-396, 1978.

[12] A.K. Bidani, K.A. Griffin, G.A. Williamson, X. Wang, and R. Loutzenhiser, "Protective importance of the myogenic response in the renal circulation," *Hypertension*, vol. 54, pp. 393-398, 2009.

[13] A.K. Bidani, A.J. Polichnowski, H. Licea-Vargas, J. Long, S. Kliethermes, G.A. Williamson, and K.A. Griffin, "BP fluctuations and the real-time dynamics of renal blood flow responses in conscious rats," *J Am Soc Nephrol*, 2019.

[14] K.A. Griffin, R. Hacioglu, I. Abu-Amarah, R. Loutzenhiser, G.A. Williamson, A.K. Bidani, "Effects of calcium channel blockers on 'dynamic' and 'steady-state step' renal autoregulation, *Am J Physiol Renal Physiol*, vol. 286, pp. F1136F1143, 2004.

[15] B. Zhao, H. Lu, S. Chen, J. Liu and D. Wu, "Convolutional neural networks for time series classification," *J Syst Eng Electron*, vol. 28, no. 1, pp. 162-169, 2017.

[16] S.R. Joshi, D.B. Headley, K.C. Ho, D. Paré and S.S. Nair, "Classification of brainwaves using convolutional neural network," in *Proc. 2019 27th European Signal Processing Conference (EUSIPCO)*, A Coruna, Spain, 2019, pp. 1-5.

[17] S. Matsui, N. Inoue, Y. Akagi, G. Nagino and K. Shinoda, "User adaptation of convolutional neural network for human activity recognition," in *Proc. 2017 25th European Signal Processing Conference (EUSIPCO)*, Kos, Greece, 2017, pp. 753-757.

[18] A.K. Bidani, K.A. Griffin, M. Picken, D.M. Lansky, "Continuous telemetric blood pressure monitoring and glomerular injury in the rat remnant kidney model," *Am J Physiol Renal Fluid Electrolyte Physiol*, vol. 265, pp. F391F398, 1993.

[19] K. Griffin, M. Picken, A. Bidani, "Radiotelemetric BP monitoring, antihypertensives and glomeruloprotection in remnant kidney model," *Kidney Int*, vol. 46, pp. 10101018, 1994.

[20] A.K. Bidani, K.D. Mitchell, M.M. Schwartz, L.G. Navar, E.J. Lewis, "Absence of progressive glomerular injury in a normotensive rat remnant kidney model, *Kidney Int*, vol. 38, pp. 2838, 1990.

[21] K.A. Griffin, A.K. Bidani, J. Ouyang, V. Ellis, M. Churchill, P.C. Churchill, "Role of endothelium-derived nitric oxide in the hemodynamic adaptations after graded renal mass reduction," *Am J Physiol Regul Integr Comp Physiol*, vol. 264, pp. R1254R1259, 1993.

[22] K.A. Griffin, M. Picken, A.K. Bidani, "Method of renal mass reduction is a critical modulator of subsequent hypertension and glomerular injury," *J Am Soc Nephrol*, vol. 4, pp. 20232031, 1994.

[23] K.A. Griffin, M.M. Picken, M. Churchill, P. Churchill, A.K. Bidani, "Functional and structural correlates of glomerulosclerosis after renal mass reduction in the rat, " *J Am Soc Nephrol*, vol. 11, pp. 497506, 2000.

[24] D.P. Kingma, J.L. Ba, "Adam: a method for stochastic optimization," in *Proc 3rd Int Conf Learning Representations*, San Diego CA, 2015.

[25] M. Fernández-Delgado, E. Cernadas, S. Barro, "Do we need hundreds of classifiers to solve real world classification problems?", *J Mach Learn Res*, vol. 15, pp. 3133-3181, 2014.