Tuijin Jishu/Journal of Propulsion Technology
ISSN: 1001-4055
Vol. 45 No. 4 (2024)

# A Novel Approach for Diagnosing Chronic Kidney Disease Using a Non-Invasive ECG Mechanism and Deep Learning Technique

**[1]K. Chapa, [2]B. Ravi**

*1Research Scholar,Department of Computer Science and Engineering,*

*GITAM (Deemed to be University), Visakhapatnam, Andhra Pradesh, India.*

*2Professor,Department of Computer Science and Engineering,*

*GITAM (Deemed to be University), Visakhapatnam, Andhra Pradesh, India.*

**Abstract**

Chronic kidney disease is one among extreme causes of increasing mortality rate. Usually, CKD diagnosis and the level of CKD identification is based on traditional serum creatinine concentration or GFR levels. Although they are reliable researchers are finding out Non-invasive methods so that the diagnostic process may become simple. ECG based CKD diagnosis is one among the non-invasive techniques upon which reasonable research is going on. Obviously, ECG is used in screening cardiovascular diseases and at the same time using the same ECG if we are able to identify the CKD it shall be more prolific to the public. In this paper, a novel mechanism is proposed in identifying the CKD level using ECG. A standard and well processed ECG dataset containing data of 10,646 subjects is taken for modelling the mechanism and computing a score which is used in turn in identifying the CKD level. A deep learning model has been used to concrete the process with which an accuracy of 99% has been obtained. Further to validate the proposed mechanism of score generation and CKD level identification, the process is applied on three datasets and obtained the prediction accuracy levels of 88.7%, 93% and 91.5% for the respective datasets which indicates the usage of ECG is quiet acceptable for CKD prediction.

**Keywords:** ECG based CKD diagnosis, Chronic Kidney Disease diagnosis, non-invasive techniques in CKD prediction

## 1. Introduction:

Chronic kidney disease is an irreversible kidney disorder which is of serious concern and also it impacts the raise of wide range of ailments that include heart failure, orthopaedic disorders, anaemia, and others. Usually, kidney disease couldn't be clearly identified at the early stage. The state of kidney becomes chronic by the time the symptoms are visible. Other diseases also lead to Chronic kidney ailments. As the chronic condition shall be known at later stage, the disease will be in an irreversible state consequently treatment becomes difficult that leads to death. Therefore, an early-stage diagnosis is highly essential for treating such ailments.

The presence of CKD is confirmed with Glomerular Filtration Rate (GFR) of 60 ml/min/1.73 m2 and/ or the indication of renal impediments for minimum three months [1]. As per the clinical settings, the blood samples of an individual are tested for the presence of Urinary albumin-creatinine ratio (UACR) and estimated GFR (eGFR) which is based on blood serum creatinine concentration. According to the KDIGO guidelines, there are five stages for renal disease in which End-stage kidney disease (ESKD) and Severe CKD are the severe stages that requires kidney replacement or yet times lead to death. The next set of stages namely the moderate CKD or mild CKD in which the CKD is curable if the patients are properly treated. The first stage is Normal CKD where the kidney disfunction is in its initial stage. Planned and prompt management increases the quality of life in CKD patients even those who possess a high risk of ESKD. Morbidity, Mortality, and healthcare related costs also may be reduced.

**Electro Cardiogram (ECG) and its Analysis:**

ECG stands for Electro Cardiogram. It is a graph representing voltage vs time in which the electrical activities of depolarization and repolarization of a cardiac muscle are characterized. The ECG graph of a normal heart beat is as shown in figure 1. It comprises a P-wave that presents atrial depolarization, a T-wave denoting the ventricular repolarization and a QRS complex representing the ventricular depolarization. PR, ST, and QT intervals are the other portions of the ECG signal.

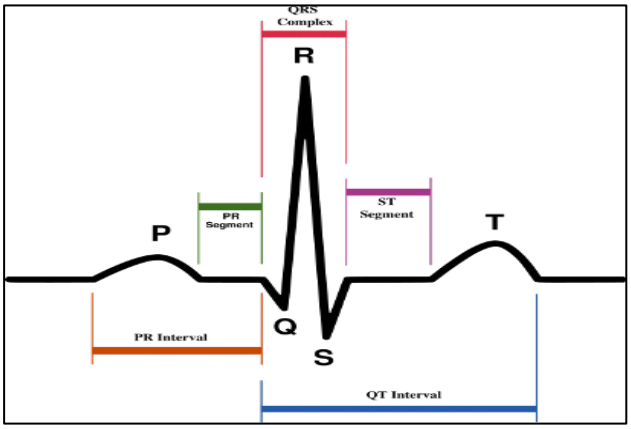



**Fig 1: Normal heart beat ECG wave form [19]**

CKD Patients may also suffer with variable proportions of cardio-vascular risk factors namely hypertension and diabetes. Also, subclinical cardio-vascular changes like myocardial fibrosis, diastolic dysfunction and left ventricular hypertrophy may take place which reflect as ECG abnormalities such as decreased T-wave amplitudes, large-amplitude T-waves, as well prolonged QRS duration in hyperkalaemia, and QTc prolongation in hypocalcaemia. There are few economical or non-invasive methods for diagnosing CKD with conventional risk calculations instead of requiring both serum and urine tests [2]. Electrocardiograms (ECGs) are low-priced, rapid, non-invasive, commonly available diagnostic tests which are frequently done during routine visits. Also, it is a quiet common test for patients with the risk of cardio related diseases. As per the current diagnostic practices, cardiologists examine ECG data and accordingly implements the treatment plans. The advent of Deep learning made the analysis of ECG waveforms easier and offering substantial performance in prediction outcomes [3]. The readings of recognizing occult arrhythmias [5][6], ventricular dysfunction [4], anaemia and age if available and DLA could be applied to screening ECGs, it could be useful potentially in CKD diagnosis. Nowadays there is a high demand for wider screening procedures and ECG enabled wearable devices are one amongst.

The ECG monitoring device captures signals from the leads and prints the signals in a wave format on a special paper called as ECG recording paper. 12 leads will be connected to the human body namely – Lead I, II, III, aVR, aVL, aVF, V1, V2, V3, V4, V5, V6.

| Monitoring cable connections | | |
|---|---|---|
| Europe | Connect to: | U.S.A. |
| Red | Right Arm | White |
| Yellow | Left Arm | Black |
| Green | Left Leg | Red |
| Black | Right Leg | Green |
| White | Chest | Brown |
| **Individual chest leads** | | |
| White / Red | C1 / V1 | Brown / Red |
| White / Yellow | C2 / V2 | Brown / Yellow |
| White / Green | C3 / V3 | Brown / Green |
| White / Brown | C4 / V4 | Brown / Blue |
| White / Black | C5 / V5 | Brown / Orange |
| White / Violet | C6 / V6 | Brown / Purple |

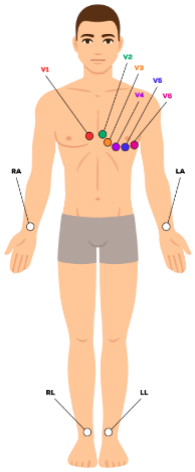



Fig 2: ECG 12 lead connections

A standard wave form printed on ECG paper is shown. Here the printing rate plays a vital role and it is generally 25 mm/sec. The ECG paper consists of grid boxes (small and large). A smaller box corresponds to 0.04 sec (40 ms) whereas a larger box includes small five boxes representing 0.20 sec (200 ms).

Vertically, height of the box represents the amplitude. Standard calibration is 10 mm corresponding to 10 small boxes which is equal to 1 mV. However double standard as well half standard considerations vary. P-wave, Q-wave, R-wave, S-wave and T-wave are the wave forms registered. Depending on the waves formed, PR segment, PR interval, QRS complex, ST segment, QT interval are characterized.

In this section a brief introduction to Chronic Kidney Disease and Electro Cardio Gram has been given while in the succeeding section the literature on how ECG is chosen as an alternative tool for CKD diagnosis is presented. The material used and the methodology implemented to reach the objective of this research is discussed in the third section. In fourth section the results were discussed. Also, the validation of the proposed mechanism is presented here. Finally, the conclusion is given in the fifth section.

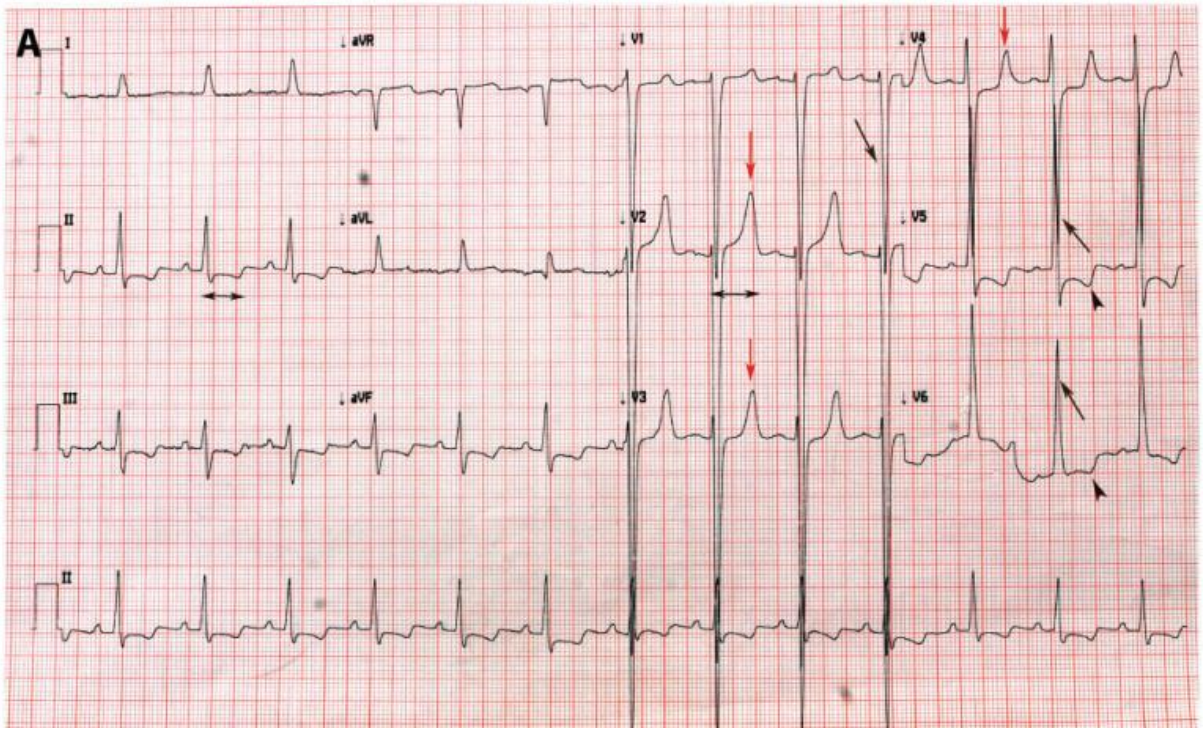



Fig 3: Standard wave form printed on ECG paper

**2. Literature study:**

Distinguished researchers made their contributions in the aspect of ECG based CKD diagnosis. The literature pertaining to their research is presented in this section.

Pannu et.al. [7] The characterization of an ECG for CKD can be manifested with the triad of findings a) narrow-based tall T waves signifying hyperkalaemia, b) prolonged QTc (>450 ms) signifying hypocalcaemia and c) high QRS voltages in chest leads ($S_{V1}$ + $R_{V5}$ or $R_{V6}$ >35 mm) with LV 'strain' (ST-segment depression and T wave inversion in left-sided leads) indicating left ventricular hypertrophy. Due to this characterization, ECG could be proposed as an important diagnostic and predictive tool for patients with kidney disorders.

Undiagnosed chronic kidney disease (CKD) is a common disorder that causes a lot of sickness and early deaths around the world. In order to identify Chronic Kidney Disease (CKD) from the collected electrocardiograms (ECGs), Lauri et al. [8] developed an advanced deep learning algorithm. The cohort contains 247,655 electrocardiograms (ECGs) pertaining to 111,370 patients between 2005 and 2019. The data was utilized to build, train, validate, and assess a deep learning model that can predict whether a patient has CKD or not from the ECG diagnosis. An additional cohort of 312,145 patients and 896,620 electrocardiograms (ECGs) recorded between 2005 and 2018 from a separate healthcare system was used to validate the model. Using 12-lead electrocardiogram (ECG) waveforms, the deep learning algorithm is able to identify various stages of chronic kidney disease (CKD) effectively. CKD detection with an AUC of 0.767, for external cohort, it achieved an AUC of 0.709. In identifying patients with CKD whose age is under 60, the model is incredibly accurate across all stages. The 12-lead and 1-lead electrocardiogram (ECG) waveforms are both used to accomplish this and the AUCs are 0.843 and 0.824 respectively. Convolutional Neural Network (CNN) is the deep learning algorithm used and it accurately detected chronic kidney disease (CKD) from electrocardiogram (ECG) waveforms.

The researchers Julario et.al. [9] in building a robust diagnostic tool, focused on determining the common clinical and laboratory variables those are predictive towards ECG abnormalities. The research enrolled 198 subjects suffering from CKD and amongst them 101 are men. Fragmented QRS and prolonged QTc were the most common ECG abnormalities found in their study. Among hospitalised patients with chronic kidney disease (CKD), it is

predicted to have a peaked T wave and prolonged QTc correlating the signs of elevated serum creatinine and haemoglobin levels. In patients with chronic kidney disease (CKD), they opine that systolic blood pressure may be a marker for a prolonged QTc and fragmented QRS.

Using electrocardiography (ECG), Kwon et.al. [10] developed and validated an easy and robust deep learning model (DLM). In their study, 115,361 patients were admitted who had previously undergone an electrocardiogram (ECG) and an eGFR measurement within 30 minutes after the first ECG. In addition, 37,190 ECGs (corresponding to 37,190 patients of another hospital) were used for external validation. The deep learning model (DLM) trained on a 12-lead electrocardiogram (ELG) was able to detect respiratory insufficiency (RI) with an area under the curve (AUC) of 0.858 (95% CI 0.851-0.866), and when tested externally, it achieved an AUC of 0.906 (0.900-0.912). Two convolution layers, two batch normalisations, one maxpooling layer, and one recurrent dropout layer made the six-stage residual block. It also makes use of two fully connected 1D layers.

Weijie et.al. [11] in their research, used a dataset consisting of over 2 million electrocardiograms (ECGs), over a thousand medical conditions, and over 250,000 patients. Their goal is to find a wide range of diseases that can be correctly diagnosed using the patient's first ECG. A total of 128 diseases and 68 disease categories have been accurately identified by the deep learning models. ResNet is the DLM used.

In [12], Badawy et.al. as a part of their research, introduced an AI-based transfer learning framework. Xception, MobileNet, MobileNetV2, MobileNetV3Large, NASNetMobile, VGG16, VGG19 and other such pre-trained convolutional neural network (CNN) models were used in this study to implement transfer learning. One among the two datasets used had four separate categories: cyst, normal, stone, and tumour. The experimental results proved that the suggested framework is better than other sophisticated classification models. With the suggested framework, an 100% success rate for five classes and a 99.98% success rate for four classes could be attained.

The research by Binsawad et.al. [13] used Optimised Forest (Opt-Forest) model. Their research possibly associates renal function and electrocardiogram (ECG) data in which the proposed OptForest's accuracy is 78.68%, 0.641 KS, and low RMSE of 0.174 demonstrates its robustness in predicting CKD compared to other ML models. This study also demonstrates the usefulness of electrocardiogram data in enhancing the early detection of chronic kidney disease.

Singh et al. [14] made their study on one hundred CKD patients. Blood urea, serum creatinine, urine, electrocardiogram, and echocardiography were the diagnostic tools used to assess the patients. In their study, electrocardiography was able to identify cardiovascular abnormalities in 72 percent of patients. One third of the patients had LVH. 68% of the total, had abnormalities detected by echocardiography. 46% of the patients had left ventricular hypertrophy. It was found that Left ventricular hypertrophy to be the common morphological abnormality. They expressed that in diagnosing left ventricular dysfunction, echocardiography serves as an extremely sensitive diagnostic tool.

Urtnasan et al. [15] in their research focused on monitoring hyperkalaemia which is diagnosed by measuring potassium levels in the serum done through blood test. But they indicate that a non-invasive alternative method for real-time monitoring of hyperkalaemia is very much required. In this method, a single-lead electrocardiogram (ECG) is analysed using a deep learning model. The study involved 29,58 patients who had episodes of hyperkalaemia between 2009 and 2019. A deep learning model that could represent the complex and repetitive pattern of cardiac activity was used by the proposed non-invasive screening tool. Using an electrocardiogram (ECG) signal, the deep learning model can noninvasively screen serum potassium levels. The deep learning model was tested on lead I, lead II, and leads V1–V6 separately. Using a single-lead electrocardiogram (ECG), the suggested non-invasive screening tool achieved remarkable results, with F1 scores of 100% on the training set, 96% on the validation set, and 95% on the test set. Therefore it could be understood that non-invasive screening is one useful tool in such diagnosis.

Galoway et al. [16] in their research focused on non-invasive procedures using the electrocardiogram (ECG) by implementing a state-of-the-art deep-learning model which is able to detect the dangerous medical condition

known as hyperkalaemia. The researchers worked out on the data of 1,576,581 electrocardiograms of 449380 patients between 1994 and 2017. The ECG data was used to train a DNN which detected Hyperkalaemia in renal disease patients with 2 ECG leads. An AUC between 0.853 and 0.883 was attained by the model.

Shafi et.al. [17] The study included 124 patients with chronic kidney disease (CKD) admitted to the nephrology ward within a 6-month period, regardless of age, who had never received renal replacement therapy before. Electrocardiograms (ECGs) with 12 leads were performed on all patients. Among the considered subjects, approximately seventy-nine percent of CKD patients have abnormalities in their electrocardiogram (ECG). Prolonged QRS duration (19.2%), tachycardia (17.6%), left and right atrial enlargement (17.6%), ST segment elevation or depression (23.4%), Q waves (27.2%), and left ventricular hypertrophy (40%) are the abnormalities.

Xu et al. [18] focused on the ESRD patients and found out that the prevalence of hyperkalaemia is said to be high among them that too it is high especially in the patients those on dialysis. They presume that traditional laboratory testing of serum potassium levels is laborious and an alternative method is essential. Using the ECG by implementing Machine Learning techniques the prevalence could be diagnosed. A total of 1024 datasets (from 2020 to 2021) that contain ECG readings were used and the data was analysed using a variety of machine learning techniques. Hyperkalaemia was predicted using 48 features from chest leads V2-V5 using a variety of machine learning models, such as Logistic Regression (LR), Support Vector Machines (SVM), Convolutional Neural Networks (CNN), Extreme Gradient Boosting (XGB), and Adaboost. Accuracy, sensitivity, specificity, F1 score, and area under the curve (AUC) are the metrics used to evaluate and compare the models' performance. Finally they supposed to say that non-invasive procedures are useful in predicting CKD quickly and accurately by analysing specific ECG waveforms using machine learning methods. In the experiments conducted, SVM performed better in the predictions, while XGB generally had a higher Area Under the Curve (AUC) in mild cases.

Having gone through the literature by all the illustrious researchers, it can be understood that in identifying CKD related and also other diseases, non-invasive techniques in association with the Artificial Intelligence, Machine and Deep Learning based approaches are the currently adopted ones. Hence in this article also, an effort is made by proposing a non- invasive ECG based CKD identification approach applied on a standard dataset, authenticated using a deep learning model and adjudicated using three datasets.

## 3. Material and Methods:

In this section, the dataset used, ECG features necessary for CKD, modelling parameters, the process of score calculation, standardization of the score process with deep learning model have been discussed.

### 3.1 Dataset

In order to diagnose a wide range of diseases, long-term electrocardiogram monitoring is a vital procedure which could be adopted and through ECG diagnosis the diseases could be detected. In a joint effort, Chapman University and Shaoxing People's Hospital (affiliated with Zhejiang University School of Medicine) created a database that contains 12-lead electrocardiogram signals for research purposes. One among the major reasons in creating this database is to help researchers with future studies on heart conditions like arrhythmia and other conditions such as CKD. Automated diagnostics that are accurate can be achieved by training modern machine learning and statistical tools on large datasets of high quality. Consequently, the researchers are able to made this standard and important dataset [19] and made available. It contains 10,646 patient ECGs recorded at 500 Hz. There are 12-dimensional and the duration of ECGs is10-seconds. It contains the labels that show the subjects' rhythms and other conditions. In order to train the algorithms, classification methods require massive datasets that cover all common types of conditions. There are more subjects, a higher sampling rate, and more leads in this dataset. Furthermore, it includes eleven cardiac rhythms and fifty-six cardiovascular disorders that are classified by specialist doctors. Furthermore, the database includes essential electrocardiogram (ECG) metrics like QRS counts, ventricle beat rate, atrial beat rate, Q offset, and T offset. The attributes of ECG dataset are as shown in figure 4.

| Attributes | Type | Value Range | Description |
|---|---|---|---|
| FileName | String | | ECG data file name (unique ID) |
| Rhythm | String | | Rhythm Label |
| Beat | String | | Other conditions Label |
| PatientAge | Numeric | 0–999 | Age |
| Gender | String | MALE/FEMAL | Gender |
| VentricularRate | Numeric | 0–999 | Ventricular rate in BPM |
| AtrialRate | Numeric | 0–999 | Atrial rate in BPM |
| QRSDuration | Numeric | 0–999 | QRS duration in msec |
| QTInterval | Numeric | 0–999 | QT interval in msec |
| QTCorrected | Numeric | 0–999 | Corrected QT interval in msec |
| RAxis | Numeric | −179~180 | R axis |
| TAxis | Numeric | −179~181 | T axis |
| QRSCount | Numeric | 0–254 | QRS count |
| QOnset | Numeric | 16 Bit Unsigned | Q onset (In samples) |
| QOffset | Numeric | 17 Bit Unsigned | Q offset (In samples) |
| TOffset | Numeric | 18 Bit Unsigned | T offset (In samples) |

**Fig 4: Attributes in ECG dataset [19]**

The dataset formulation procedure comprises four steps. First, all the subjects had undergone a 12 lead ECG for 10 seconds. A medical professional made the diagnosis of irregular heartbeat and other cardiac issues. A separate licenced medical professional carried out a secondary check. Also, a particular naming convention was employed to export ECG data and diagnostic information into XML files using the GE (General Electric) MUSE system. The conversion tool can finally take an XML file containing electrocardiogram (ECG) data and other diagnostic information as input and converts it to a CSV file. Later a sequential noise reduction technique was applied to examine and improve raw ECG data. To remove signals above 50 Hz, a Butterworth low pass filter was used, since the typical electrocardiogram (ECG) frequency range is 0.5 Hz to 50 Hz. The impact of baseline wandering was mitigated by using an LOESS smoother. To tackle the residual noise, the Non Local Means (NLM) method was finally used.

The two requirements laid out by the standard ECG measurement protocol was met by the dataset. Lead II's voltage reading must always equal the sum of lead I and lead III's voltage readings; secondly, the sum of the voltage readings from leads $a_{VR}$, $a_{VL}$, and $a_{VF}$ must be zero; No change will be made to the corresponding ECG data if the operator moves the right hand electrode to the left hand or vice versa. Furthermore, a flat line on the ECGs might be the result of electrodes coming loose during the test. An algorithm that detects unwanted instances checks for errors automatically. Then the database of the ECG records is made clean. Finally a standard dataset is said to be offered by the authors.

### 3.2 ECG measures required for CKD Analysis observed in ECG:

For CKD analysis, adoption of ECG makes the process of simple and beneficial to any individual. The ECG depicts CKD prevalence as shown in figure 5. The red arrows represent narrow-based tall T waves that signify hyperkalemia whereas the double horizontal arrows represent the prolonged QTc (>450 ms) signifying hypocalcemia.

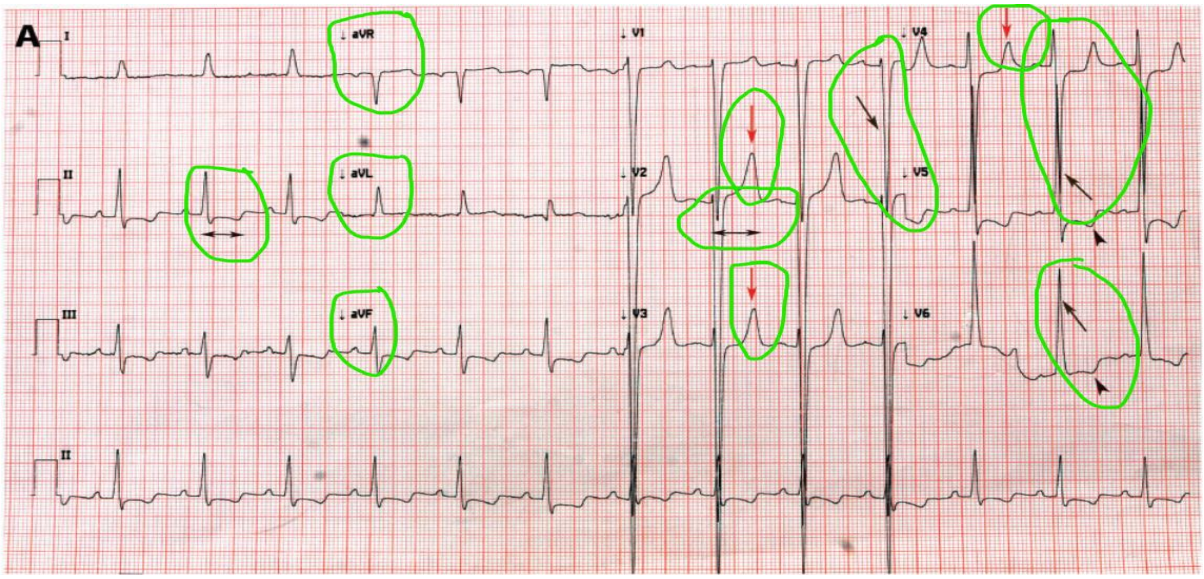



**Fig 5: ECG indicating the abnormal features of CKD**

On the other hand, high black arrows represent the QRS voltages in chest leads ($S_{V1}$ + $R_{V5}$ or $R_{V6}$ >35 mm) with LV 'strain' while the black arrowheads represents the ST-segment depression and T wave inversion in left-sided leads indicating the left ventricular hypertrophy. Finally, the trio of prolonged QTc, left ventricular hypertrophy and tall T waves designate the advanced stage requiring renal replacement therapy or otherwise said as ESRD.

### 3.3 Considerable Electrocardiographic characteristics according to chronic kidney disease (CKD) stage:

As mentioned in the earlier subsection, CKD could be identified observing the ECG graph. Also, it could be done through various readings of T-wave, QRS and so on. The features with which the CKD could be predicted and their ranges are as shown in figure 6. The list of ECG diagnostic features in the dataset and the considered features for CKD score generation mechanism are as mentioned in the table 1.

| Electrocardiographic characteristic | Overall | No CKD | Mild CKD | Moderate CKD | ESRD | P value |
|---|---|---|---|---|---|---|
| Heart rate, bpm | 81.1 ±21.8 | 80.5 ±22.1 | 85.7 ±21.2 | 82.8 ±21.3 | 80.9 ±19.2 | <0.001 |
| PR interval, ms | 166.0 ±37.7 | 164.5 ±36.9 | 165.8 ±36.8 | 172.5 ±43.2 | 168.4 ±34.6 | <0.001 |
| P wave duration, ms | 52.2 ±9.6 | 52.1 ±9.2 | 52.1 ±11.2 | 51.9 ±11.1 | 53.6 ±9.2 | <0.001 |
| QRS duration, ms | 101.0 ±29.0 | 98.7 ±27.2 | 112.6 ±36.3 | 111.8 ±35.6 | 98.9 ±24.7 | <0.001 |
| QTc interval, ms | 457.6 ±66.7 | 452.6 ±70.3 | 477.4 ±60.2 | 474.5 ±55.2 | 468.3 ±43.7 | <0.001 |
| P-wave axis | 50.2 ±27.5 | 50.3 ±27.0 | 50.6 ±26.6 | 50.2 ±29.9 | 48.9 ±26.2 | <0.001 |
| R-wave axis | 23.6 ±60.2 | 23.2 ±51.1 | 38.0 ±85.1 | 24.6 ±75.1 | 22.1 ±57.2 | <0.001 |
| T-wave axis | 57.5 ±57.6 | 54.1 ±54.9 | 61.1 ±64.4 | 67.6 ±66.3 | 67.0 ±57.3 | <0.001 |

| **ECG diagnostics available:** | **Considered features for CKD score:** |
|---|---|
| FileName | Rhythm (Rh) |
| Rhythm | PatientAge (PA) |
| Beat | QTCorrected (QTC) |
| PatientAge | Raxis (R) |
| Gender | Taxis (T) |
| VentricularRate | QRSCount (QRS) |
| AtrialRate | Gender |
| QRSDuration | |
| QTInterval | |
| QTCorrected | |
| RAxis | |
| TAxis | |
| QRSCount | |
| QOnset | |
| QOffset | |
| TOffset | |

**Fig 6: ECG Features representing CKD**

**Table 1: Available features in dataset and features considered for the proposed method.**

In the above figure the characteristics of ECG modelling are mentioned as Heartrate, P,T,R wave values, PR interval, QRS and QTC values. However, in the prior discussion, it can be understood that to characterize ECG for CKD, the PR interval and P wave values are not much significant. Hence only the most important features are considered in the proposed computation. For each parameter considered to model the mechanism, the ranges and respective score are mentioned in tables 2(a) to 2(d) in the next subsection.

Although as per KDIGO guidelines, the stages are 1, 2, 3a, 3b, 4 and 5; in our mechanism, the considered stages are four. NO CKD, Mild CKD, Moderate CKD and Severe CKD. The stages 1,2 are represented as Mild CKD, 3a & 3b as Moderate CKD while 4 and 5 stages are considered as Severe CKD.

**3.4 CKD Score generation parameters and value ranges:**

| QRS | | QTC | | R | |
|---|---|---|---|---|---|
| **Range** | **Score** | **Range** | **Score** | **Range** | **Score** |
| 80-100 | 1 | 350-460 | 1 | 30-100 | 1 |
| 70-80 or 100-110 | 2 | 330-350 or 460-480 | 2 | 20-30 or 100-110 | 2 |
| 60-70 or 110-120 | 3 | 310-330 or 480-500 | 3 | 10-20 or 110-120 | 3 |
| <60 or >120 | 4 | <310 or >500 | 4 | <10 or >120 | 4 |

**Table 2(a) : QRS, QTC, R parameters and their ranges**

| T | | Rhythm | | Patient Age | |
|---|---|---|---|---|---|
| **Range** | **Score** | **Type** | **Score** | **Range** | **Score** |
| **-15 - 105** | **1** | **SR, SB** | **1** | **10-30** | **1** |
| **-25 : -15 or 105-115** | **2** | **SAWWR** | **2** | **30-50** | **2** |
| **-35 : -25 or 115-125** | **3** | **AFIB, SI, AT, ST** | **3** | **50-70** | **3** |
| **<-35 or >125** | **4** | **AVRT, AF, AVNRT, SVT** | **4** | **>70** | **4** |

**Table 2(b) : T, Rhythm, Patient Age parameters and their ranges**

| Gender | |
|---|---|
| **Female** | **1** |
| **Male** | **2** |

**Table 2(c) : Gender parameter**

| CKD Score | |
|---|---|
| **Range** | **CKD level** |
| **7 - 10** | **No CKD (1)** |
| **11 - 16** | **Mild CKD (2)** |
| **17 - 22** | **Moderate CKD (3)** |

| CKD Score calculation (Example) | | |
|---|---|---|
| **Feature** | **Value** | **Score** |
| **QRSC** | **114** | **3** |
| **QTC** | **496** | **3** |
| **R** | **81** | **1** |
| **T** | **27** | **1** |
| **Rhy** | **AFIB** | **3** |
| **Age** | **85** | **4** |
| **Gender** | **M** | **2** |

| | |
|---|---|
| 23 - 26 | ESRD (4) |

**Table 2(d) : Final CKD score range and CKD level.**

| | | |
|---|---|---|
| Total Score & CKD level | 17 | 3 |

**Table 2(e) : CKD level calculation example**

Considering all the parameters and their ranges, an example computation for a record in the chosen dataset is presented in table 2(e) above. As the QRSC value is 114 its score is 3 referring to QRSC parameter table. Similarly computing the individual scores of all parameters and adding them the obtained value is 17 with which the CKD level is designated as 3. This computation is performed for all the records in the dataset and included as another attribute in the original ECG dataset. Now the dataset consists of CKD level for each ECG diagnostic record.

### 3.4 Proposed mechanism - Algorithm for ECG-CKD score generation for entire ECG dataset using Deep Learning:

In this section, the algorithm for Score generation and identifying the CKD level for the records in the chosen dataset is presented. Also the procedure is outlined in the figure 7. The ECGCKD_Score_main() contains the major steps of the entire procedure. The number of subjects are 10646 ($S_1$ to $S_{10646}$) i.e, the number of records of individuals in the dataset. The number of features is seven i.e., $F_1$ to $F_7$. Considering each subject and respective features, the score computation is done as per the procedure mentioned in ECKD_Compute() method. Then the level of CKD is computed using CKD_Compute() method.

**ECKD_Compute($ECG_{ij}$, *i, j):***

*j = 1:* if(80<= $ECG_{ij}$ <=100)

$ECKD_{ij}$ ← 1

elseif (70<= $ECG_{ij}$<80) and (100< $ECG_{ij}$<=110)

$ECKD_{ij}$ ← 2

elseif (60<= $ECG_{ij}$<70) and (110< $ECG_{ij}$<=120)

$ECKD_{ij}$ ← 3

else

$ECKD_{ij}$ ← 4

*j = 2:* if(350<= $ECG_{ij}$ <=460)

$ECKD_{ij}$ ← 1

elseif (330<= $ECG_{ij}$<350) and (460< $ECG_{ij}$<=480)

$ECKD_{ij}$ ← 2

elseif (310<= $ECG_{ij}$<330) and (480< $ECG_{ij}$<=500)

$ECKD_{ij}$ ← 3

else

$ECKD_{ij}$ ← 4

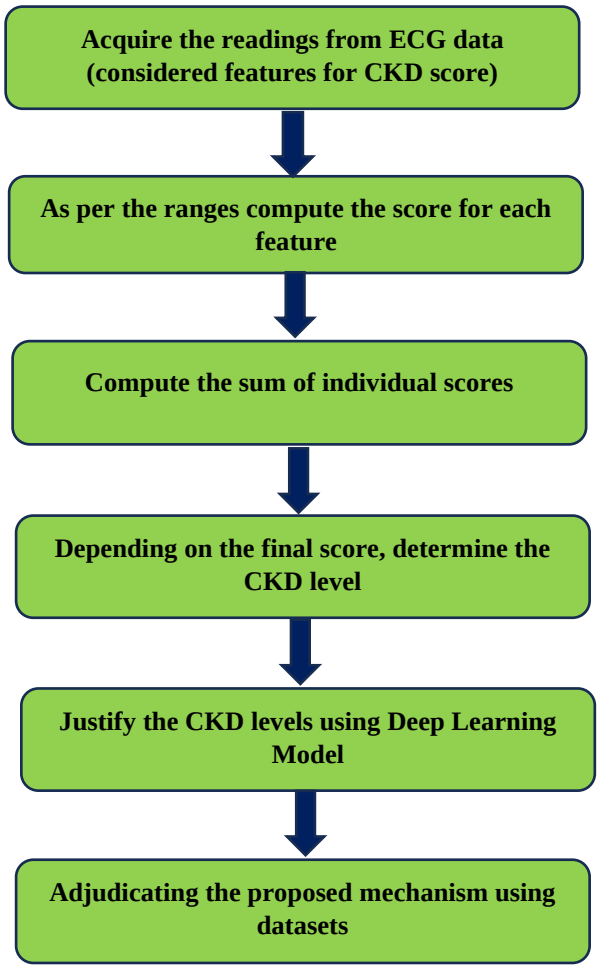



Fig 7: Proposed ECG based CKD diagnosis method

*j = 3:* if(30<= $ECG_{ij}$ <=100)

$ECKD_{ij}$ ← 1

elseif (20<= $ECG_{ij}$<30) and (100< $ECG_{ij}$<=110)

$ECKD_{ij}$ ← 2

elseif (10<= $ECG_{ij}$<20) and (110< $ECG_{ij}$<=120)

$ECKD_{ij}$ ← 3

else

$ECKD_{ij}$ ← 4

*j = 4:* if(-15<= $ECG_{ij}$ <=105)

$ECKD_{ij}$ ← 1

elseif (-25<= $ECG_{ij}$<15) and (105< $ECG_{ij}$<=115)

$ECKD_{ij}$ ← 2

elseif (-35<= $ECG_{ij}$<-25) and (115< $ECG_{ij}$<=125)

$ECKD_{ij}$ ← 3

else

$ECKD_{ij}$ ← 4

*j = 5:* if($ECG_{ij}$ equals ‘SR’ or ‘SB’)

$ECKD_{ij}$ ← 1

elseif ($ECG_{ij}$ equals ‘SAWWR’)

$ECKD_{ij}$ ← 2

elseif ($ECG_{ij}$ equals ‘AFIB’ or ‘SI’ or ‘ST’ or ‘AT’)

$ECKD_{ij}$ ← 3

elseif ($ECG_{ij}$ equals ‘AVRT’ or ‘AF’ or ‘AVNRT’ or ‘SVT’)

$ECKD_{ij}$ ← 4

*j = 6:* if(10<= $ECG_{ij}$ <=30)

$ECKD_{ij}$ ← 1

elseif (30< $ECG_{ij}$ <=50)

$ECKD_{ij}$ ← 2

elseif (50< $ECG_{ij}$ <=70)

$ECKD_{ij}$ ← 3

else

$ECKD_{ij}$ ← 4

*j = 7:* if($ECG_{ij}$ equals 'FEMALE')

$ECKD_{ij}$ ← 1

elseif ($ECG_{ij}$ equals 'MALE')

$ECKD_{ij}$ ← 2

**CKD_Compute(*ECKD_Score$_i$*):**

if(7<= $ECKD\_Score_i$ <=10)

$CKD\_Level_i$ ← 1

elseif (11<= $ECKD\_Score_i$ <=16)

$CKD\_Level_i$ ← 2

elseif (17<= $ECKD\_Score_i$ <=22)

$CKD\_Level_i$ ← 3

elseif (23<= $ECKD\_Score_i$ <=26)

$CKD\_Level_i$ ← 4

**ECGCKD_Score_main():**

*S* represents Subject, *F* represents Feature

$ECG_{ij}$ represents the value of $F_j$ for $S_i$

$ECKD_{ij}$ represents the score of Feature *j* for Subject *i*

Step 1: $F_1$ ← *QRS*, $F_2$ ← *QTC*, $F_3$ ← *Raxis*, $F_4$ ← *Taxis*, $F_5$ ← *Rh*, $F_6$ ← *Age*, $F_7$ ← *Gender*

Step 2: for each *i* and *j* read $ECG_{ij}$

Step 3: for each *i* and *j*, $ECKD_{ij}$ = ECKD_Compute($ECG_{ij}$, *i, j)*

Step 4: for each Subject $S_i$,

$ECKD\_Score_i$ ← $\sum_{i,j} ECKD_{ij}$

$CKD\_Level_i$ ← CKD_Compute($ECKD\_Score_i$)

Step 5: for p in 1 to 0.8*count(ECG_data) $ECG\_Train_p$ ← random_select(ECG_data)

Step 6: ECG_Test ← ECG_data - ECG_Train

Step 7: DLModel = <IL,DL1,DL2,DL3,OL,Sigmoid,CCE,Adam,softmax>

Step 8: DLModel_train(ECG_Train)

Step 9: DLModel_test(ECG_Test)

Step 10: Accuracy = DLModel_accuracy

Now the dataset contains all the features which are available originally and in addition the CKD Level also is appended. Further to make the dataset justifiable, a Deep Learning based Neural Network procedure is

implemented. As a final step, validation of the proposed mechanism is done using three real world datasets. The validate procedure is as below.

**Validate_ECGCKD(Dataset)**

Step 1: Read_Dataset()

Step 2: ECKD_Compute() for all features and subjects in dataset

Step 3: CKD_Compute for all subjects

Step 4: for each subject compare computed CKD_Level and Dataset CKD_Level

Step 5: Identify the number of matches

Step 6: Compute Validation_Accuracy()

$$\text{Validation\_Accuracy} = \frac{\text{CKD\_Level}_i\text{Match\_Count}}{\text{CKD\_Level}_i\text{Total\_Count}} * 100$$

**4. Results and Discussion:**

In this section, firstly the deep learning model used for the CKD score generation, CKD level prediction and the associated results are discussed. Then the validation process of the proposed CKD score mechanism for three datasets along with the respective results are discussed.

**4.1 Deep Learning Model and the associated results:**

The deep learning model for the justification of applying the CKD score generation and prediction mechanism is as shown in figure 8. The model is built using 5 layers namely one input layer, three hidden layers and one output layer.

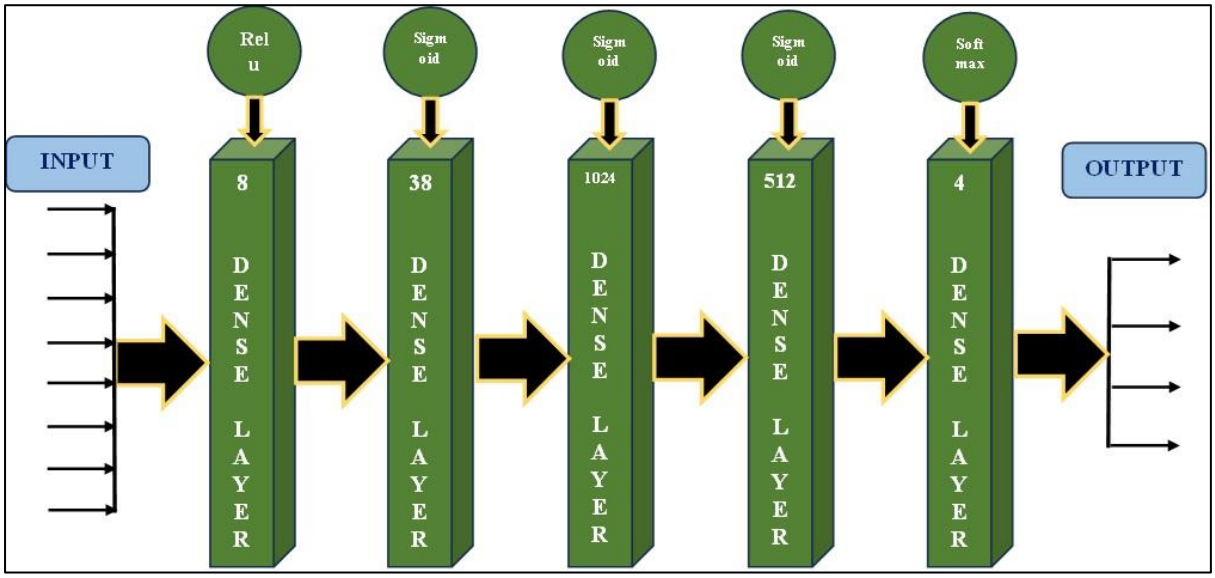



**Fig 8: Deep Learning model architecture**

All the layers including input and output layers are dense in nature. The number of input features are 8 and outputs are 4 levels of CKD. The activation function used at the input layer is relu while for the other layers it is sigmoid function. The optimizer used is Adam and the loss function is Cross Conditional Entropy. Finally at the output layer softmax classifier is used. As all the hyper parameters used are suitable and sufficient the classification accuracy obtained is high. The dataset is divided into train data containing 8517 (80%) records and test data containing 2129 (20%) records. The network is trained using the train data and tested on the test data towards which the classification has been done perfectly. The training and testing accuracies are 100% and 99.3%

respectively. Hence the proposed mechanism is justifiable in nature. The total number of parameters used is 567,194 and are as specified below.

```
Model: "sequential_1"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 dense_5 (Dense)             (None, 8)                 64

 dense_6 (Dense)             (None, 38)                342

 dense_7 (Dense)             (None, 1024)              39936

 dense_8 (Dense)             (None, 512)               524800

 dense_9 (Dense)             (None, 4)                 2052

=================================================================
Total params: 567,194
Trainable params: 567,194
```

In addition to the above, to authenticate the proposed mechanism, it is subjected to validation using three real world datasets which is described in the next section.

**4.2 Validating the proposed mechanism by applying on other datasets:**

**4.2.1 Dataset 1 paper [20]:**

In this study the authors worked on the data of 191 patients. Among these patients nearly 92.1% are having ECG abnormalities in one or other form. QTc interval, peaked T wave, left ventricular hypertrophy, fragmented QRS complex, poor R wave progression are the most common abnormalities found in ECG in 36.6%, 22%, 16.7%,29.8% and 24.6% respectively. The number of patients CKD level wise is not given but altogether the CKD prevalence is seen in 176 patients. The proposed mechanism diagnosed 169 patients with 88.4% classification. The classification comparison is as shown in table 3 and figure 9.

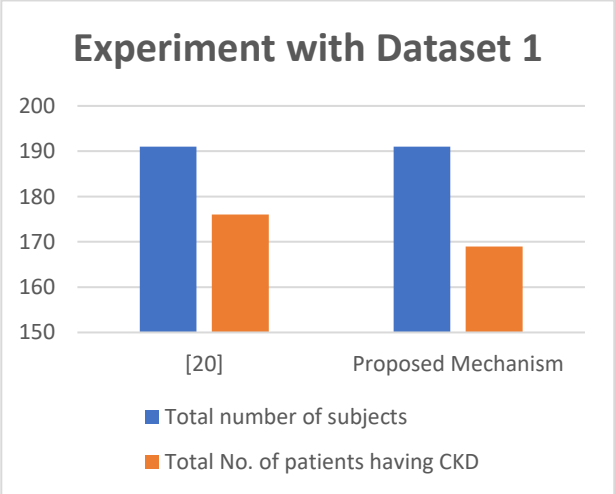



Fig 9: Classification comparison

| Methodology | Total number of subjects | Total No. of patients having CKD | Percentage of classification |
|---|---|---|---|
| **[20]** | 191 | 176 | 92.1 |
| **Proposed Mechanism** | 191 | 169 | 88.47 |

Table 3: classification comparison

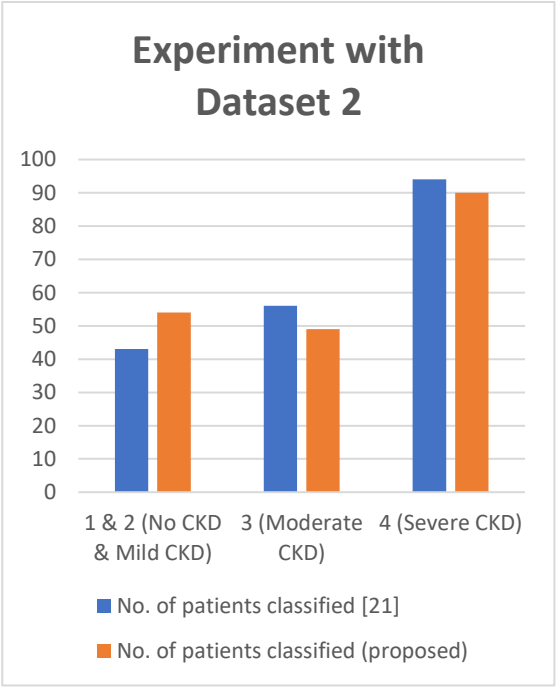
Moreover, the CKD level wise patients as classified by the proposed mechanism are mentioned in table 4.

| CKD level | No. of patients | % |
|---|---|---|
| **1 (No CKD)** | 22 | 11.5 |
| **2 (Mild CKD)** | 45 | 23.5 |
| **3 (Moderate CKD)** | 68 | 35.6 |
| **4 (Severe CKD)** | 58 | 30.3 |

**Table 4: CKD level wise patients classification with dataset 1**

**4.2.2 Dataset 2 paper [21]:**

The total number of subjects considered in this dataset is 193. The distribution of patients is done in five stages namely More Severe (Stage 5), Severe (Stage 4), Moderate (Stage 3), Mild to Moderate (Stage 2) and Mild (Stage 1) while in the experiment done using proposed mechanism the more severe and Severe stages are considered as Severe CKD (level 4), Moderate stage and Mild to Moderate stage are considered as Moderate CKD (level 3). Mild CKD (Stage 1) and those patients having no CKD are considered as No CKD (level 1& 2) in our proposed mechanism. The Number of patients classified and the variation in percentage between the proposed mechanism and earlier implemented mechanism are presented in table 5 and figure 10. As per our proposed mechanism among the 193 patients, 180 patients are suffering with CKD while remaining 13 patients are not having CKD. Fig 10: Classification comparison




| CKD level | No. of patients classified [21] | No. of patients classified (proposed) | Variation in % |
|---|---|---|---|
| **1 (No CKD)** | 43 | 13 | 5.7 |
| **2 (Mild CKD)** | | 41 | |
| **3 (Moderate CKD)** | 56 | 49 | 3.6 |
| **4 (Severe CKD)** | 94 | 90 | 2 |

**Table 5: CKD level wise classification performance with dataset 2**

**4.2.3 Dataset 3 paper [22]:**

In this study 190 subjects were considered for study to compare specific changes in ECG as markers of various arrhythmias in CKD patients as well patients with ESRD. Among them the number of ESRD patients is 75 on regular hemodialysis whereas the number of patients with stage 3–5 CKD is 75 and 40 healthy subjects. The ESRD patients are considered as severe CKD patients (level 4), CKD levels 3-5 patients are considered as Moderate CKD patients (level 3) and the healthy patients are considered as Level 1 and 2. The Number of patients classified and the variation in percentage between the proposed mechanism and earlier implemented mechanism are presented in table 6 and figure 11.

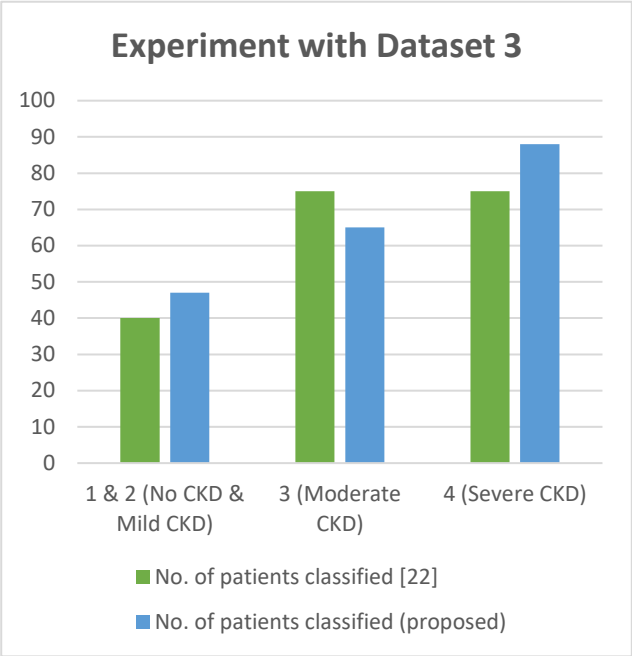



Fig 11: Classification comparison

As per the proposed mechanism among the 190 patients, 174 patients are suffering with CKD while remaining 16 patients are not having CKD.

| CKD level | No. of patients classified [22] | No. of patients classified (proposed) | Variation in % |
|---|---|---|---|
| **1 (No CKD)** | 40 | 16 | 3.68 |
| **2 (Mild CKD)** | | 31 | |
| **3 (Moderate CKD)** | 75 | 65 | 5.26 |
| **4 (Severe CKD)** | 75 | 88 | 6.84 |

**Table 6: CKD level wise classification performance with dataset 3**

In the three datasets, the number of subjects considered are 191, 193 and 190 respectively. In these datasets the number of patients suffered with CKD are 169, 180 and 174 and the respective percentages of classification are 88.47, 93.26 and 91.57. In consolidation, the classification percentage and the average differences in classification are shown in table 7.

| | No. of subjects considered | CKD prevalence % Classification – proposed method | Average % Variation with Original method |
|---|---|---|---|
| **Dataset 1 [20]** | 191 | 88.47 | 3.63 |
| **Dataset 2 [21]** | 193 | 93.26 | 3.76 |
| **Dataset 3 [22]** | 190 | 91.57 | 5.26 |

**Table 7: CKD classification performance by proposed method and average variation % with original method in the 3 datasets**

From the above table it could be understood that the prevalence of CKD is around 90% in all three cases i.e., with three datasets. Also, the average percentage of variation of classification between the original methods and the proposed method is around 3 – 5 % which is a less value. Hence the proposed mechanism of ECG based CKD diagnosis is said to be validated and can be used in identifying CKD.

**5. Conclusion:**

Chronic Kidney Disease is a life-threatening disease and if it is not treated properly or if it is neglected, the severity leads to death. Once any individual is effected with kidney disease and not properly treated, it silently deprives the functionality of kidney functionality progressively and at some time leads to kidney failure. In general blood serum tests are used in diagnosing CKD. Due to the inconvenience faced by the patients and few other reasons non-invasive techniques are also gaining attention. In this research work, ECG based CKD diagnosis is considered to be an alternative means. An ECG dataset is considered for the proposed method of generating CKD score using essential parameters and the respective CKD Levels are generated. Not only the CKD levels are identified, they are also justified using a deep learning model with 99.3% accuracy. Finally, the proposed mechanism is adjudicated using three datasets containing 191, 193 and 190 patient records respectively. In the three cases considered the proposed mechanism arrived at classifying CKD prevalence with respective percentages of 88.47, 93.26 and 91.57 respectively. Moreover, the average differences (in percentages) of CKD prevalence classification in the three datasets correspondingly are 3.63, 3.76 and 5.26. As the classification percentages are around 90% and the percentage differences are less, the proposed mechanism could be considered as an alternative means for CKD diagnosis.

**References:**

[1] Gudeti, Bhavya & Mishra, Shashvi & Malik, Shaveta & Fernandez, Terrance Frederick & Tyagi, Amit & Kumari, Shabnam. (2020). A Novel Approach to Predict Chronic Kidney Disease using Machine Learning Algorithms. 1630-1635. 10.1109/ICECA49313.2020.9297392.

[2] Bang, H. et al. SCreening for Occult REnal Disease (SCORED): a simple prediction model for chronic kidney disease. Arch Intern Med 167, 374–381 (2007).

[3] Raghunath, S. et al. Prediction of mortality from 12-lead electrocardiogram voltage data using a deep neural network. Nat. Med. 26, 886–891 (2020).

[4] Attia, Z. I. et al. Screening for cardiac contractile dysfunction using an artificial intelligence-enabled electrocardiogram. Nat. Med. 25, 70–74 (2019).

[5] Kwon, J. M. et al. A deep learning algorithm to detect anaemia with ECGs: a retrospective, multicentre study. Lancet Digit Health 2, e358–e367 (2020).

[6] Attia, Z. I. et al. An artificial intelligence-enabled ECG algorithm for the identification of patients with atrial fibrillation during sinus rhythm: a retrospective analysis of outcome prediction. Lancet. 394, 861–867 (2019)

[7] A K Pannu, M Dutta, H M Singh, ECG triad of chronic kidney disease, *QJM: An International Journal of Medicine*, Volume 112, Issue 6, June 2019, Pages 465–466, https://doi.org/10.1093/qjmed/hcy302

[8] Lauri Holmstrom, Matthew Christensen, Neal Yuan, J. Weston Hughes, John Theurer, Melvin Jujjavarapu, Pedram Fatehi, Alan Kwan, Roopinder K. Sandhu, Joseph Ebinger, Susan Cheng, James Zou, Sumeet S. Chugh and David Ouyang, "Deep learning based electrocardiographic screening for chronic kidney disease", (2022),https://doi.org/10.1101/2022.03.01.22271473

[9] Rerdin Julario, Ricardo Adrian Nugraha, Bagus Putra Dharma Khrisna, Tony Santoso Putra, Eka Prasetya Budi Mulia, Ryan Enast Intan, Firas Farisi Alkaff medRxiv 2021.01.15.21249916; doi: https://doi.org/10.1101/2021.01.15.21249916

[10] Kwon JM, Kim KH, Jo YY, Jung MS, Cho YH, Shin JH, Lee YJ, Ban JH, Lee SY, Park J, Oh BH. Artificial intelligence assessment for early detection and prediction of renal impairment using electrocardiography. Int Urol Nephrol. 2022 Oct;54(10):2733-2744. doi: 10.1007/s11255-022-03165-w. Epub 2022 Apr 11. PMID: 35403974; PMCID: PMC9463260.

[11] Sun, Weijie & Kalmady, Sunil & Salimi, Amir & Sepehrvand, Nariman & Ly, Eric & Hindle, Abram & Greiner, Russ & Kaul, Padma. (2022). ECG for high-throughput screening of multiple diseases: Proof-of-concept using multi-diagnosis deep learning from population-based datasets. 10.48550/arXiv.2210.06291.

[12] Badawy M, Almars AM, Balaha HM, Shehata M, Qaraad M and Elhosseini M (2023) A two-stage renal disease classification based on transfer learning with hyperparameters optimization. Front. Med. 10:1106717. doi: 10.3389/fmed.2023.1106717

[13] Muhammad Binsawad, Enhancing kidney disease prediction with optimized forest and ECG signals data,Heliyon, Volume 10, Issue 10, 2024, e30792,ISSN 2405-8440, https://doi.org/10.1016/j.heliyon.2024.e30792.

[14] Dr. K.D. Singh, Dr. D.K. Malviya and Dr. Anchal Upadhyay, "Study of ECG and Echocardiographic Changes in Patient with Chronic Kidney Disease (CKD)", (2022), ISSN 2515-8260, European Journal of Molecular & Clinical Medicine, Vol(9)(4) pp 552-559

[15] Urtnasan E, Lee JH, Moon B, Lee HY, Lee K, Youk H, Noninvasive Screening Tool for Hyperkalemia Using a Single-Lead Electrocardiogram and Deep Learning: Development and Usability Study, JMIR Med Inform 2022;10(6):e34724, doi: 10.2196/34724

[16] Galloway CD, Valys AV, Shreibati JB, et al. Development and Validation of a Deep-Learning Model to Screen for Hyperkalemia From the Electrocardiogram. *JAMA Cardiol*. 2019;4(5):428–436. doi:10.1001/jamacardio.2019.0640

[17] Shafi S, Saleem M, Anjum R, Abdullah W, Shafi T. ECG Abnormalities In Patients With Chronic Kidney Disease. J Ayub Med Coll Abbottabad. 2017 Jan-Mar;29(1):61-64. PMID: 28712176.

[18] Daojun Xu, Bin Zhou, Jiaqi Zhang, Chenyu Li, Chen Guan, Yuxuan Liu, Lin Li, Haina Li, Li Cui, Lingyu Xu, Hang Liu, Li Zhen & Yan Xu (2023) Prediction of hyperkalemia in ESRD patients by identification of multiple leads and multiple features on ECG, Renal Failure, 45:1, 2212800, DOI: 10.1080/0886022X.2023.2212800

[19] Zheng, Jianwei; Rakovski, Cyril; Danioko, Sidy; Zhang, Jianming; Yao, Hai; Hangyuan, Guo (2019). A 12-lead electrocardiogram database for arrhythmia research covering more than 10,000 patients. figshare. Collection. https://doi.org/10.6084/m9.figshare.c.4560497

[20] Mulia, Eka P.B., Nugraha, Ricardo A., A'yun, Maya Q., Juwita, Rahima R., Yofrido, Filipus M., Julario, Rerdin and Alkaff, Firas F.. "Electrocardiographic abnormalities among late-stage non-dialysis chronic kidney disease patients" *Journal of Basic and Clinical Physiology and Pharmacology*, vol. 32, no. 3, 2021, pp. 155-162. https://doi.org/10.1515/jbcpp-2020-0068

[21] Akshat Jain, Geetanshu Goel, Sujeet Kamtalwar. Study of electrocardiographic changes and their comparison in different stages of chronic kidney disease. MedPulse International Journal of Medicine. September 2019; 11(3): 193-197. https://www.medpulse.in/Medicine/

[22] Yehia, H., Youssef, G., Gamil, M. *et al.* Electrocardiographic substrates of arrhythmias in patients with end-stage and chronic kidney diseases: a case–control study. *Egypt Heart J* **75**, 13 (2023). https://doi.org/10.1186/s43044-023-00338-5